AD-A008 205

A BIVARIATE TEST OF GOODNESS OF FIT BASED ON A GRADUALLY INCREASING NUMBER OF ORDER STATISTICS

Jose Kreimerman

Cornell University

Prepared for:

Army Research Office
Office of Naval Research

March 1975

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #250	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD-A008 205
4. TITLE (and Subtitle) A BIVARIATE TEST OF GOODNESS OF FIT BASED ON A GRADUALLY INCREASING NUMBER OF ORDER STATISTICS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) José Kreimerman		8. CONTRACT OR GRANT NUMBER(s) DAHC04-73-C-0008 N00014-67-A-0077-0020
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research College of Engineering, Cornell University Ithaca, New York 14853		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Sponsoring Military Activity U.S. Army Research Office Durham, N.C. 27706		12. REPORT DATE March 1975
		13. NUMBER OF PAGES ~~80~~ 85
11. CONTROLLING OFFICE NAME AND ADDRESS Sponsoring Military Activity Statistics and Probability Program Office of Naval Research Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

tests of fit
asymptotic distribution theory
order statistics
multivariate distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report is concerned with the problem of testing the goodness of fit of a sample drawn from a continuous bivariate distribution, where the null hypothesis is that the true distribution is a completely specified one. A test is developed, where the test criterion is based on some standardized order statistics, standardization which depends on the distribution under the null hypothesis.

By order statistics from a bivariate distribution we will understand the following. We order the sample by the first coordinate and we select some

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

of the ordered first coordinates as strip limits. For the points in the interior of each strip, we order their second coordinate and we consider some of those.

The distribution theory, of which the test of goodness of fit is an application, is presented in a general framework. The sample is drawn from a bivariate continuous distribution which may depend on the sample size n. We prove the asymptotic joint normality of an increasing number of standardized order statistics, standardization which depends on the true distribution. We assume that as n increases, we gradually increase the number of strips; that in each strip the number of those second coordinates we consider is also gradually increased; that we move slowly into the tails; that the derivative of the marginal probability density function exists and its absolute value is bounded by a number depending on n; that the derivative of the conditional marginal probability density function for the second coordinate within each strip, conditional on the strip limits, exists and its absolute value is bounded by a number depending on n; that there exist bounds, which depend on n, for both marginal p.d.f.s mentioned above; that there are some asymptotic relationships concerned with the speed of moving into the tails, the slowly increasing number of standardized order statistics, the separation in order of those order statistics and the smoothness of both marginal p.d.f.s. Using a generalization of the concept of limiting distribution, asymptotic equivalence, which allows us to handle the case where the dimension of the multivariate random variables depends on n, we prove that those standardized order statistics are asymptotically equivalent to a multivariate normal with mean zero and known covariance matrix.

Using this result, a test of goodness of fit is developed, and the test criterion is the quadratic form of this multivariate normal. It follows that under the null hypothesis, its distribution approximates the chi-square distribution with known degrees of freedom. In a particular case, it is shown that under the alternative the test criterion has approximately a noncentral chi-square distribution with known noncentrality parameter and known degrees of freedom.

DEPARTMENT OF OPERATIONS RESEARCH
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

TECHNICAL REPORT NO. 250

March 1975

A BIVARIATE TEST OF GOODNESS OF FIT BASED ON A GRADUALLY INCREASING NUMBER OF ORDER STATISTICS

by

José Kreimerman

Prepared under contracts

DAHC04-73-C-0008, U. S. Army Research Office - Durham

and

N00014-67-A-0077-0020, Office of Naval Research

ib

TABLE OF CONTENTS

Chapter I

INTRODUCTION

1.1. Introduction

An important area of Statistical Inference relates to the problem of assessing the conformity or goodness of fit of some observations to a null hypothesis. This hypothesis may be that the observations came from a distribution which belongs to a given family or may be a completely specified distribution. A statistical procedure which solves such a problem is called a test of goodness of fit.

This thesis is concerned with the problem of testing the goodness of fit of a sample drawn from a continuous bivariate distribution, where the null hypothesis is that the true distribution is a completely specified one. A test is developed, where the test criterion is based in some functions of a subset of order statistics, functions which depend on the distribution under the null hypothesis.

By order statistics from a bivariate distribution we will understand the following. We order the sample by the first coordinate and we select some of the ordered first coordinates as strip limits. For the points in the interior of each strip, we order their second coordinate and will consider some of those. Our test criterion is based on some functions of the strip limits and on functions of those second coordinates ordered within each strip.

The distribution theory on which this test is based will show the asymptotic joint normality of those functions of the subset of order statistics on which the test criterion is based on, assuming that as

the sample size increases, we gradually increase the number of strips, that in each strip the number of those ordered second coordinate we consider is also gradually increased, that we move slowly into the tails for both coordinates, that for each sample size n, the derivative of the marginal probability density function (p.d.f.) exists and its absolute value is bounded by a number depending on n, that the derivative of the conditional marginal p.d.f. for the second coordinate within each strip, conditional on the strip limits, exists and its absolute value is bounded by a number depending on n, that for each n, there exist bounds, which depend on n, for both marginals mentioned above. This distribution theory allows the distribution from which we sample to depend on n. Those order statistics taken into consideration, are standardized and the standardization depends on the true distribution.

Using a generalization of the concept of limiting distribution, asymptotic equivalence, which allows us to handle the case where the dimension of the multivariate random variables depends on n, we will prove that those functions of the subset of order statistics are asymptotically equivalent to a multivariate normal with mean zero and known coveriance matrix.

As the test criterion is the quadratic form of this multivariate normal it follows that under the null hypothesis, its distribution approaches the chi-square distribution with degrees of freedom known and depending on the number of strips and the number of order statistics considered within the strips.

1.2. Summary

In Section 1.1 we have given an overview of the problem and of the results obtained, and in Section 1.3 we give a review on the historical background. In this Section we outline briefly the contents of this thesis.

In Chapter II we study the asymptotic distribution of a gradually increasing number of some functions of order statistics from a continuous bivariate distribution.

In Section 2.1 we will introduce the concept of order statistics from a bivariate distribution, and for each n, we will define some functions of them, functions which depend on the distribution. For each n, we will define a multivariate random variable, whose dimensionality increases with n and we will also define a multivariate normal variable with the same dimensionality as the other one for each n.

In Section 2.2 we will state the assumptions that will hold throughout Chapter II and deduce from them some results.

In Section 2.3 we will explicitly show the joint probability density function of both the functions of the order statistics and the multivariate normal, and we will state some results useful for the proofs of Section 2.4.

In Section 2.4 we will prove that both sequences of multivariate random variable, in one case the functions of the order statistics and in the other the multivariate normal, are asymptotically equivalent.

In Chapter III, we are going to develop a bivariate test of goodness of fit, for testing the hypothesis that the sample comes from a completely specified continuous bivariate distribution.

In Section 3.1 we propose the test of goodness of fit, whose test criterion is based on the functions of order statistics defined in Section 2.1.

In Section 3.2 we show that in a particular case the random variable used as test criterion has under the alternative approximately a noncentral chi-square distribution with noncentrality parameter known and known degrees of freedom. We also give a concrete example, which satisfies all the assumptions given in Section 2.2.

In Appendix A we present the concept of asymptotic equivalence of sequences of multivariate random variables, whose dimensionality depends on the position in the sequence. The historical background is given in Section A.1 and the concept and some basic results are presented in Section A.2.

The notation used in Sections 1.3 and A.1 are independent of that used in the rest of the thesis, and the notation we adopt in this thesis will be introduced as it is needed in the development.

1.3. Historical Background

Historically, the first clear use of a test of goodness of fit seems to have been by Bernoulli [2], regarding the closeness of orbital planes of the planets to one another and to the equatorial plane of the sun.

1.3.1. The Chi-Square Test of Goodness of Fit

Karl Pearson [16], was the first one to give a proposal for the general use of such a procedure in a broad class of situations, namely

his chi-square test of goodness of fit, where the essential features of this type of procedure were presented without ambiguity. In this pioneering paper he introduced his X^2-test criterion and he realized that this criterion, under certain circumstances follows the chi-square distribution. Pearson essentially operated as follows: He divided the observations into k-cells, defined x_i as the number of observations which lie in cell i, let m_i be the expected number of observations falling in cell i and defined $X^2 = \sum_{i=1}^{k} \frac{(x_i - m_i)^2}{m_i}$. He first dealt with the case where m_i $(1 \leq i \leq k)$ are known numbers. He assumed that the x_i may be taken as normally distributed, so implicitly he was committed to the assumption that the expectations m_i are large for all cells, and he proved that if the null hypothesis is true X^2 has asymptotically a chi-square distribution with $(k-1)$ degrees of freedom. The test rejects the null hypothesis if X^2 is too large.

Pearson established the necessary distribution theory for finding significance levels when the null hypothesis provides the exact values for the m_i $(1 \leq i < k)$, but he did not show that the exact distribution of X^2, always discontinuous, actually approaches chi-square as a limiting distribution. A fully rigorous proof may be found in Cramér [4].

Fisher [6], generalized the procedure for the problem when the expected number of observations lying in each cell are estimated from the sample.

The exact distribution for X^2, for fixed sample size, is usually extremely complicated, and for this reason most of the literature concentrates on generalizing the procedure for large sample size.

Following those early papers in the area, there is an extensive literature trying to improve different aspects of the test. Yates [19], under certain conditions, introduced a correction for continuity. Neyman [15], showed that if we test the null hypothesis that the expectations are m_i when in fact they are m_i' and if m_i, m_i' and the significance level are kept fixed, then as n increases, the power of the test tends to 1.

When X^2 is used to test the hypothesis that the observations came from a continuous distribution, we must, in order to apply the test, decide in advance how to group the observations into cells. Both the number of cells and the division points between cells are at our disposal, and the choices we make will affect the sensitivity of the test. Mann and Wald [13] and Gumbel [8] suggested choosing the cells such that each cell has the same expected number, namely n/k. Other authors, recently suggested using randomly assigned cell limits.

Nevertheless, in testing the hypothesis that the sample came from a given continuous distribution versus the alternative that it came from a continuous distribution belonging to a given family, the power of the test is very poor for an alternative where each test cell has the same expected number under both the null hypothesis and this alternative, no matter how large the sample size.

This is not the unique criticism that the traditional form of the chi-square test has undergone, but we only mention this particular one because the test based on a gradually increasing number of order statistic proposed by Weiss [22] meets this objection and we are generalizing his results for the bivariate case.

1.3.2. The Kolmogorov-Smirnov Test of Goodness of Fit

Kolmogorov [12] and Smirnov [17] have suggested a test of goodness of fit. A brief description of the test for the univariate case would be the following:

Let $X_1, X_2, \ldots, X_n$ be random variables drawn from an unknown cumulative distribution function $F(x)$, let $S_n(x)$ be the observed cumulative step function of the sample, (i.e., $S_n(x) = \frac{k(x)}{n}$ where $k(x)$ is the number of observations less or equal to x), called the empirical c.d.f., then the sampling distribution of $D_n = \sup_x |F(x) - S_n(x)|$ is known, and is independent of $F(x)$, if $F(x)$ is continuous.

Therefore, a natural test of goodness of fit for testing the hypothesis that $F(x) = F_0(x)$, a continuous specified c.d.f., is to reject the hypothesis if $\sup_x |F_0(x) - S_n(x)| > d_n(\alpha)$, where $d_n(\alpha)$ is taken from a table.

Kolmogorov-Smirnov proved that $\sqrt{n}D_n$ has a limiting distribution as n increases. A table for the limiting distribution was given by Smirnov [18].

Although the criticism we were applying to the chi-square test does not apply to the Kolmogorov-Smirnov, unfortunately when we are interested in testing goodness of fit for a continuous bivariate distribution, D_n's distribution depends on the true distribution of the observation, even when the null hypothesis is true (i.e., it is not distribution-free any more) and actually its distribution is difficult to compute.

1.3.3. Weiss Test of Goodness of Fit

Weiss [22] developed a test of goodness of fit, for the problem of testing the hypothesis that the sample comes from a continuous univariate distribution completely specified. His test criterion is a quadratic form of some standardized order statistics, standardization which depends on the true distribution.

The test is based on the asymptotic normality of a gradually increasing number of order statistics. The number of order statistics he considered, depends on the sample size n. He assumed the existence of the derivative of the p.d.f. and that the absolute value is bounded by a n depended number. That as n increases, he moves slowly into the tails. That the p.d.f.'s are bounded from above and below by numbers depending on n. And that there exist some asymptotic conditions between all these elements, and between them and the number of order statistics which are between those considered. The standardization is made with the population quantile, its p.d.f. value and it is assumed that the ordered statistic is the corresponded sample quantile. It is also assumed that the number of order statistics considered increases as n increases.

As the test criterion is the quadratic form valued at those standardized order statistics, under the null hypothesis, its distribution approximates the chi-square distribution with as many degrees of freedom as number of order statistics considered.

Chapter II

ON THE ASYMPTOTIC JOINT NORMALITY OF A GRADUALLY INCREASING NUMBER OF SOME FUNCTIONS OF ORDER STATISTICS FROM A CONTINUOUS BIVARIATE DISTRIBUTION

2.0. Introduction

This chapter is concerned with the asymptotic joint distribution of a gradually increasing number of order statistics from a continuous bivariate distribution.

In Section 2.1 we will introduce the concept of order statistics from a bivariate population, and for each sample size n, we will define some functions of them, functions which depend on the distribution. Then for each n, we will define some multivariate random variable, whose dimensionality increases with n. We will also define some multivariate normal variable, with the same dimension as the one depending on the order statistics.

In Section 2.2 we will state the assumptions that will hold throughout this chapter and deduce from them some useful results.

In Section 2.3 we will explicitly show the joint probability density function of both the functions of the order statistics and the multivariate normal, and we will state some elementary results, which will be used in the proofs of the following section.

In Section 2.4 we will formulate and prove a theorem showing that both multivariate random variables, the function of the order statistics and the multivariate normal are asymptotically equivalent

(for details see Appendix A), or in other words, their distributions are asymptotically approaching each other.

2.1. Order Statistics from a Bivariate Distribution, and a Multivariate Normal Distribution

In this section we will define, for each positive integer n, two multivariate random variables. One is a set of functions, which depend on the distribution, of some order statistics from a sample of size n drawn from a continuous bivariate distribution; and the other is a multivariate normal with mean zero and some known covariance matrix.

2.1.1. Order Statistics from a Bivariate Distribution

By order statistics from a bivariate distribution we will understand the following. We order the sample by the first coordinate and we select some of the ordered first coordinates as strip limits. For the points in the interior of the strip, we order their second coordinate and will consider some of those.

Let $(X_{1i}(n), X_{2i}(n))$ $(1 \le i \le n)$ be independent identically distributed (i.i.d.) continuous bivariate random variables with common probability density function (p.d.f.) $f(x_1, x_2, n)$ and common cumulative distribution function (c.d.f.) $F(x_1, x_2, n)$. Let $Y_1(1) < \ldots < Y_1(n)$ be the ordered $X_{1i}(n)$'s. For each n, let p_1, q_1, $0 < p_1 < q_1 < 1$ be such that np_1 and nq_1 are positive integers, and let K_1 and L_1 be positive integers, such that

$$K_1 L_1 = n(q_1 - p_1) . \tag{2.1}$$

Let

$$m = L_1 - 1 . \tag{2.2}$$

Let $S(j,n) = \{(X_{1i}(n),X_{2i}(n));\ Y_1(np_1+(j-1)L_1)<X_{1i}(n)<Y_1(np_1+jL_1)$, $(1\leq i\leq n)\}$ $(1\leq j\leq K_1)$ and let $Y_2(1,j)<\ldots<Y_2(m,j)$ $(1\leq j\leq K_1)$ be the ordered values of the second coordinate of the elements of $S(j,n)$. For each n , let p_2 , q_2 , $0<p_2<q_2<1$ such that mp_2 and mq_2 are positive integers and let K_2 and L_2 be positive integers such that

$$K_2L_2 = m(q_2-p_2) . \tag{2.3}$$

Let

$$d = K_1 + 1 + K_1(K_2+1) . \tag{2.4}$$

Define the d-dimensional random variable $Y(n)$ as

$$[Y_1(np_1+(j-1)L_1) \quad (1\leq j\leq K_1+1) \ , \ Y_2(mp_2+(i-1)L_2,j) \quad (1\leq i\leq K_2+1, 1\leq j\leq K_1)]$$

the vector $Y(n)$ contains the order statistics we are interested in, and we are going to define some functions of them, which depend on the distribution.

Let $F_{1n}(x)$ and $f_{1n}(x)$ be the marginal c.d.f. and marginal p.d.f. of the X_1's , and let $G_{2n}(x,j)$ and $g_{2n}(x,j)$ $(1\leq j\leq K_1)$ be the conditional marginal c.d.f. and conditional marginal p.d.f. of the second coordinates of the elements of $S(j,n)$, conditional on $[Y_1(np_1+(j-1)L_1)$ $(1\leq j\leq K_1+1)]$. We next define

$$T_1(j) = F_{1n}^{-1}\left(\frac{np_1+(j-1)L_1}{n}\right) \quad , \quad t_1(j) = f_{1n}(T_1(j)) \qquad (1\leq j\leq K_1+1) \tag{2.5}$$

$$T_2(i,j) = G_{2n}^{-1}\left(\frac{mp_2+(i-1)L_2}{m},j\right) \ , \ t_2(i,j) = g_{2n}(T_2(i,j),j) \tag{2.6}$$
$$(1\leq i\leq K_2+1 , 1\leq j\leq K_1)$$

$$W_1(j) = \sqrt{n}\,t_1(j)[Y_1(np_1+(j-1)L_1)-T_1(j)] \qquad (1\leq j\leq K_1+1) \tag{2.7}$$

$$W_2(i,j) = \sqrt{m}\,t_2(i,j)[Y_2(mp_2+(i-1)L_2,j)-T_2(i,j)] \tag{2.8}$$
$$(1\leq i\leq K_2+1 , 1\leq j\leq K_1)$$

and $W(n)$ as a d-dimensional random variable whose coordinates are given by (2.7) and (2.8), that is

$$W(n) = [W_1(j) \quad (1\leq j\leq K_1+1) \ , \ W_2(i,j) \quad (1\leq i\leq K_2+1 , 1\leq j\leq K_1)] \tag{2.9}$$

and let $g(w(n),n)$ be its joint p.d.f. whose exact expression will be given in Section 2.3.

2.1.2. A Multivariate Normal Distribution

For each n , define the positive real numbers A_1 , A_2 , A_3 , A_4 and the symmetric matrix $V(n) = \{v_{ij}(n)\}$ of size d as follows

$$A_1 = \frac{L_1}{np_1} \quad , \quad A_2 = \frac{L_1}{n(1-q_1)} \quad , \quad A_3 = \frac{L_2}{mp_2} \quad , \quad A_4 = \frac{L_2}{m(1-q_2)}$$

$$
v_{ij}(n) = \begin{cases}
\dfrac{L_1^2[1+(i-1)A_1][1+(K_1+1-j)A_2]}{nm(K_1A_1A_2+A_1+A_2)} & (1\le i\le j\le K_1+1) \\[2ex]
\dfrac{L_2^2[1+(i'-1)A_3][1+(K_2+1-j')A_4]}{m(L_2-1)(K_2A_3A_4+A_3+A_4)} & \begin{pmatrix} i'=i-[K_1+1+(k-1)(K_2+1)] \\ j'=j-[K_1+1+(k-1)(K_2+1)] \\ 1\le k\le K_1\ ,\ 1\le i'\le j'\le K_2+1 \\ K_1+2\le i\le j\le d \end{pmatrix} \\[2ex]
0 & \text{otherwise.}
\end{cases} \tag{2.10}
$$

Define $Z(n)$ as a d-dimensional random variable, normally distributed with mean zero and covariance matrix $V(n)$.

$$Z(n) = [Z_1(j) \quad (i\le j\le K_1+1)\ ,\ Z_2(i,j) \quad (1\le i\le K_2+1\ ,\ 1\le j\le K_1)]$$

and let $h(z(n),n)$ be its joint p.d.f. whose expression will be given in Section 2.3.

2.1.3. Some Additional Notation

In this subsection we are going to define some quantities which are going to be useful throughout this thesis. Define

$$B_{11} = \text{g.l.b.}\{x;F_{1n}(x)>0\} \qquad (\text{may be } -\infty\ ) \tag{2.11}$$

$$B_{12} = \text{l.u.b.}\{x;F_{1n}(x)<1\} \qquad (\text{may be } +\infty\ ) \tag{2.12}$$

$$B_{21}(j) = \text{g.l.b.}\{x;G_{2n}(x,j)>0\} \qquad (\text{may be } -\infty\ ) \qquad (1\le j\le K_1) \tag{2.13}$$

$$B_{22}(j) = \text{l.u.b.}\{x; G_{2n}(x,j)<1\} \quad (\text{may be } +\infty) \quad (1\le j\le K_1) \tag{2.14}$$

$$b_1 = \inf_x \{f_{1n}(x); T_1(1)\le x\le T_1(K_1+1)\} \tag{2.15}$$

$$B_1 = \sup_x \{f_{1n}(x); T_1(1)\le x\le T_1(K_1+1)\} \tag{2.16}$$

$$b_2 = \min_{1\le j\le K_1} \inf_x \{g_{2n}(x,j); T_2(1,j)\le x\le T_2(K_2+1,j)\} \tag{2.17}$$

$$B_2 = \min_{1\le j\le K_1} \sup_x \{g_{2n}(x,j); T_2(1,j)\le x\le T_2(K_2+1,j)\} \tag{2.18}$$

The reader is reminded that throughout this thesis the numbers A_1, A_2, A_3, A_4, b_1, b_2, B_1, B_2, B_{11}, B_{12}, $B_{21}(j)$ $(1\le j\le K_1)$, $B_{22}(j)$ $(1\le j\le K_1)$, d, K_1, K_2, L_1, L_2, m, p_1, p_2, q_1, q_2 depend on n but this dependence is not explicitely shown for notational convenience.

2.2. Assumptions and Some Consequences

Throughout this section we are going to list the assumptions, deduce some results based directly on the assumptions, and state some properties of the coordinates of $Z(n)$.

2.2.1. Assumptions

Our assumptions are given in Table I.

Table I

Assumption	#	Assumption	#
$\lim_{n\to\infty} p_1 = 0$	(1)	For all n and all x in the open interval (B_{11}, B_{12}), $f'_{1n}(x)$ exists and $\lvert f'_{1n}(x)\rvert < D_1$ where D_1 is a real number, which depends on n, and $D_1 \geq 1$ for all n.	(12)
$\lim_{n\to\infty} q_1 = 1$	(2)	For all n and j, $(1 \leq j \leq K_1)$, and all x in the open interval $(B_{21}(j), B_{22}(j))$, $g'_{2n}(x,j)$ exists and $\lvert g'_{2n}(x,j)\rvert < D_2$, where D_2 is a real number, which depends on n and $D_2 \geq 1$ for all n.	(13)
$\lim_{n\to\infty} np_1 = \infty$	(3)	$b_1 \leq 1$ for all sufficiently large n	(14)
$\lim_{n\to\infty} n(1-q_1) = \infty$	(4)	$b_2 \leq 1$ for all sufficiently large n	(15)
$\lim_{n\to\infty} K_1 = \infty$	(5)	$\lim_{n\to\infty} \frac{L_1 D_1}{n b_1^2} = 0$	(16)
$\lim_{n\to\infty} p_2 = 0$	(6)	$\lim_{n\to\infty} \frac{B_1 K_1}{\sqrt{L_1}\, b_1} = 0$	(17)
$\lim_{n\to\infty} q_2 = 1$	(7)	$\lim_{n\to\infty} \frac{B_1 K_1 D_1}{\sqrt{n}\, b_1^3} = 0$	(18)
$\lim_{n\to\infty} \frac{K_1}{mp_2} = 0$	(8)		
$\lim_{n\to\infty} \frac{K_1}{m(1-q_2)} = 0$	(9)		
$\lim_{n\to\infty} K_2 = \infty$	(10)		
For all n, np_1, mp_2, nq_1, mq_2, K_1, K_2, L_1, L_2 are positive integers	(11)		

Table I (Continued)

Assumption	#	Assumption	#
$\lim\limits_{n\to\infty} \dfrac{p_1 D_1^2}{n b_1^4} = 0$	(19)	$\lim\limits_{n\to\infty} \dfrac{B_2 K_2 D_2}{\sqrt{m}\, b_2^3} = 0$	(25)
$\lim\limits_{n\to\infty} \dfrac{(1-q_1) D_1^2}{n b_1^4} = 0$	(20)	$\lim\limits_{n\to\infty} \dfrac{p_2 D_2^2}{m b_2^4} = 0$	(26)
$\lim\limits_{n\to\infty} \dfrac{K_1 \sqrt{D_1}}{b_1 \sqrt{L_1}} = 0$	(21)	$\lim\limits_{n\to\infty} \dfrac{(1-q_2) D_2^2}{m b_2^4} = 0$	(27)
$\lim\limits_{n\to\infty} \dfrac{n D_1}{L_1^{1.5} b_1^2} = 0$	(22)	$\lim\limits_{n\to\infty} \dfrac{K_1 K_2 \sqrt{D_2}}{\sqrt{L_2}\, b_2} = 0$	(28)
$\lim\limits_{n\to\infty} \dfrac{L_2 D_2}{m b_2^2} = 0$	(23)	$\lim\limits_{n\to\infty} \dfrac{m K_1 D_2}{L_2^{1.5} b_2^2} = 0$	(29)
$\lim\limits_{n\to\infty} \dfrac{B_2 K_2}{\sqrt{L_2}\, b_2} = 0$	(24)	$\lim\limits_{n\to\infty} \sqrt{\dfrac{n}{p_1}}\, t_1(1)[B_{11} - T_1(1)] = -\infty$	(30)

Assumption	#
$\lim\limits_{n\to\infty} \sqrt{\dfrac{n}{1-q_1}}\; t_1(K_1+1)[B_{12} - T_1(K_1+1)] = +\infty$	(31)
$\lim\limits_{n\to\infty} \max\limits_{1\le j\le K_1} \dfrac{K_1 \sqrt{p_2}}{\sqrt{m}\, b_2 \lvert B_{21}(j) - T(1,j) \rvert} = 0$	(32)
$\lim\limits_{n\to\infty} \max\limits_{1\le j\le K_1} \dfrac{K_1 \sqrt{1-q_2}}{\sqrt{m}\, b_2 [B_{22}(j) - T_2(K_2+1,j)]} = 0$	(33)

2.2.2. Some Useful Results Based on the Assumptions

Since by assumption (12) $D_1 > 1$ for all n, and by (14) $b_1 \leq 1$ for all large n, then $0 < (K_1)/(\sqrt{L_1}) \leq (K_1\sqrt{D_1})/(\sqrt{L_1}b_1)$ for large n and by (21) we have

$$\lim_{n\to\infty} \frac{K_1}{\sqrt{L_1}} = 0 \tag{2.19}$$

and $0 < n/L_1^{1.5} \leq (nD_1)/(L_1^{1.5}b_1^2)$, and by (22) we have

$$\lim_{n\to\infty} \frac{n}{L_1^{1.5}} = 0\ . \tag{2.20}$$

Since

$$0 < \frac{D_1K_1}{\sqrt{n}b_1^2} < \frac{D_1K_1}{\sqrt{L_1}b_1^2} = \frac{D_1K_1L_1}{L_1^{1.5}b_1^2} < \frac{nD_1}{L_1^{1.5}b_1^2}$$

by (2.1), then by assumption (22) we have

$$\lim_{n\to\infty} \frac{D_1K_1}{\sqrt{n}b_1^2} = 0\ . \tag{2.21}$$

We are now going to show that for the following pairs (r,s): $(1,1), (0,2), (0,3), (1,2), (2,1), (3,0)$,

$$\lim_{n\to\infty} \frac{nD_1^s}{L^{s+(r/2)}b_1^{2s}} = 0\ . \tag{2.22}$$

Since (2.22) holds for $(r,s) = (1,1)$ by assumption (22) it follows that for large n $0 < (nD_1)/(L_1^{1.5}b_1) < 1$, and therefore $0 < (nD_1^2) < (n^2D_1^2)/(L_1^3b_1^4) < (nD_1)/(L_1^{1.5}b_1^2)$ which implies that (2.22) also holds

for $(r,s) = (0,2)$. Since for large n , $0 < (nD_1^3)/(L_1^3 b_1^6) <$ $(n^3 D_1^3)/(L_1^{4.5} b_1^6) < (nD_1)/(L_1^{1.5} b_1^2)$ then (2.22) holds for $(r,s) = (0,3)$. Given that we proved that (2.22) holds for $(r,s) = (0,2)$ and $0 <$ $(nD_1^2)/(L_1^{2.5} b_1^4) < (nD_1^2)/(L_1^2 b_1^4)$, it follows that it also holds for $(r,s) = (1,2)$. By the fact that $0 < (nD_1)/(L_1^2 b_1^2) < (nD_1)/(L_1^{1.5} b_1^2)$ then by assumption (22), (2.22) holds for $(r,s) = (2,1)$, and finally (2.20) shows that (2.22) also holds for $(r,s) = (3,0)$. Therefore the proof that (2.22) holds for (r,s): $(1,1),(0,2),(0,3),(1,2),(2,1),(3,0)$ is completed.

Since by assumption (13) $D_2 > 1$ and by (15) $b_2 \leq 1$ for all large n , then for large n , $0 < (K_1 K_2)/(\sqrt{L_2}) < (K_1 K_2 \sqrt{D_2})$ and by (28)

$$\lim_{n\to\infty} \frac{K_1 K_2}{\sqrt{L_2}} = 0 \tag{2.23}$$

and for large n , $0 < (mK_1)/(L_2^{1.5}) < (mK_1 D_2)/(L_2^{1.5} b_2^2)$ and by (29)

$$\lim_{n\to\infty} \frac{mK_1}{L_2^{1.5}} = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{n}{L_2^{1.5}} = 0 . \tag{2.24}$$

Since $0 < (K_1 K_2 D_2)/(\sqrt{m} b_2^2) < (K_1 K_2 D_2)/(\sqrt{L_2} b_2^2) < (mK_1 D_2)/(L_2^{1.5} D_2)$ then by assumption (29)

$$\lim_{n\to\infty} \frac{K_1 K_2 D_2}{\sqrt{m} b_2^2} = 0 . \tag{2.25}$$

Using assumption (29) and the same kind of argument as those used to prove (2.22), it follows that

$$\lim_{n\to\infty} \frac{mK_1 D_2^s}{L^{s+(r/2)} b_2^{2s}} = 0 \tag{2.26}$$

for the pairs $(r,s) = (1,1),(0,2),(0,3),(1,2),(2,1)$ and $(3,0)$

If we would remove the conditions that for all n, $D_1 \geq 1$ and $D_2 \geq 1$ in assumptions (12) and (13) then we should include (2.19), (2.21), (2.23), (2.25) as assumptions and replace (22) by (2.22) and (29) by (2.26).

2.2.3. Some Variances Related to the Multivariate Normal

Throughout this thesis $\sigma^2(x)$ will denote the variance of x and $\sigma(x)$ the standard deviation of x.

From (2.10) and recalling (2.1) and (2.3), if assumptions (1) through (10) hold, then

$$\sigma^2(Z_1(1)) = p_1(1+\delta_1) \tag{2.27}$$

$$\sigma^2(Z_1(K_1+1)) = (1-q_1)(1+\delta_1') \tag{2.28}$$

$$\sigma^2(Z_1(j+1)-Z_1(j)) = \frac{L_1^2}{nm}\left(1 - \frac{A_1A_2}{K_1A_1A_2+A_1+A_2}\right) = \frac{L_1}{n}(1+\delta_1'') \quad (1\leq j\leq K_1) \tag{2.29}$$

$$\sigma^2(Z_2(1,j)) = p_2(1+\delta_2) \qquad (1\leq j\leq K_1) \tag{2.30}$$

$$\sigma^2(Z_2(K_2+1,j)) = (1-q_2)(1+\delta_2') \qquad (1\leq j\leq K_1) \tag{2.31}$$

$$\sigma^2(Z_2(i-1,j)-Z_2(i,j)) = \frac{L_2^2}{m(L_2-1)}\left(1 - \frac{A_3A_1}{K_2A_3A_4+A_3+A_4}\right) = \frac{L_2}{m}(1+\delta_2'') \tag{2.32}$$

$$(1 \leq i \leq K_2 \ , \ 1 \leq j \leq K_1)$$

where δ_1 , δ_1' , δ_1'' , δ_2 , δ_2' , δ_2'' depend on n and their common limit as n increases is zero.

Furthermore, since all the variances of the coordinates of $Z(n)$ approach zero as n increases, there exists a real number $\bar{\sigma}^2$ such that

$$\sigma^2(Z_1(j)) \leq \bar{\sigma}^2 \quad (1 \leq j \leq K_1+1) \quad \text{and}$$

$$\sigma^2(Z_2(i,j)) \leq \bar{\sigma}^2 \quad (1 \leq i \leq K_2+1, 1 \leq j \leq K_1) \tag{2.33}$$

2.3. The Joint P.D.F.s of the Functions of Order Statistics, The Multivariate Normal, and Other Known Facts

In Subsection 2.3.1 we will write the p.d.f.s of $W(n)$ and $Z(n)$ in a convenient way for the purpose of making it easier to follow the proofs in this chapter.

The p.d.f. for $W(n)$ is found by computing the p.d.f. of $Y(n)$ in two stages: first the marginal for the Y_1's and then the conditional marginal of the Y_2's conditioned on the Y_1's , taking advantage of the fact that the Y_2's of different strips are independent and making the change of variables.

In Subsection 2.3.2, we will state some known facts that will be useful for the proofs of Section 2.4.

2.3.1. The P.D.F.s

The joint p.d.f. for $W(n)$, $g(w(n),n)$, can be written as the product of the following twelve expressions.

$$\frac{n!}{(np_1-1)!(n-nq_1)![m!]^{K_1} n^{(K_1+1)/2}} \tag{2.34}$$

$$\frac{1}{\prod_{j=1}^{K_1+1} t_1(j)} \tag{2.35}$$

$$[F_{1n}(T_1(1) + \frac{w_1(1)}{\sqrt{n}t_1(1)})]^{np_1-1} \tag{2.36}$$

$$[1 - F_{1n}(T_1(K_1+1) + \frac{w_1(K_1+1)}{\sqrt{n}t_1(K_1+1)})]^{n-nq_1} \tag{2.37}$$

$$\prod_{j=1}^{K_1} [F_{1n}(T_1(j+1) + \frac{w_1(j+1)}{\sqrt{n}t_1(j+1)}) - F_{1n}(T_1(j) + \frac{w_1(j)}{\sqrt{n}t_1(j)})]^m \tag{2.38}$$

$$\sum_{j=1}^{K_1+1} f_{1n}(T_1(j) + \frac{w_1(j)}{\sqrt{n}b_1(j)}) \tag{2.39}$$

$$\frac{1}{\sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} t_2(i,j)} \tag{2.40}$$

$$\prod_{j=1}^{K_1} [G_{2n}(T_2(1,j) + \frac{w_2(1,j)}{\sqrt{m}t_2(1,j)}, j)]^{mp_2-1} \tag{2.41}$$

$$\prod_{j=1}^{K_1} [1 - G_{2n}(T_2(K_2+1,j) + \frac{w_2(K_2+1,j)}{\sqrt{m}t_2(K_2+1,j)}, j)]^{m-mq_2} \tag{2.42}$$

$$\prod_{j=1}^{K_1} \prod_{i=1}^{K_2} [G_{2n}(T_2(i+1,j) + \frac{w_2(i+1,j)}{\sqrt{m}t_2(i+1,j)}, j) - G_{2n}(T_2(i,j) + \frac{w_2(i,j)}{\sqrt{m}t_2(i,j)}, j)]^{L_2-1} \tag{2.43}$$

$$\prod_{j=1}^{K_1} \prod_{i=1}^{K_2+1} g_{2n}(T_2(i,j) + \frac{w_2(i,j)}{\sqrt{m}t_2(i,j)}, j) \tag{2.44}$$

$$\frac{[m!]^{K_1}}{[(mp_2-1)!(m-mq_2)![(L_2-1)!]^{K_2} m^{(K_2+2)/2}]^{K_1}} \tag{2.45}$$

if $B_{11} < T_1(1) + (w_1(1))/(\sqrt{n}t_1(1)) <\ldots< T_1(K_1+1) + (w_1(K_1+1))/(\sqrt{n}t_1(K_1+1)) < B_{12}$ and $B_{21}(j) < T_2(1,j) + (w_2(1,j))/(\sqrt{m}t_2(1,j)) <\ldots< T_2(K_2+1,j) + (w_2(K_2+1,j))/(\sqrt{m}t_2(K_2+1,j)) < B_{22}(j)$ $(1 \le j \le K_1)$ and $g(w(n),n) = 0$ otherwise.

The joint p.d.f. for $Z(n)$, $h(z(n),n)$, recalling the fact that $Z(n)$ is a d-variate normal with mean zero and covariance matrix $V(n)$, can be written as the product of the following five expressions

$$\left(\frac{nm}{L_1^2}\right)^{\frac{K_1+1}{2}} \left[\frac{1}{p_1(1-q_1)}\right]^{\frac{1}{2}} \left[\frac{L_1^2 K_1}{n^2} + \frac{(1-q_1+p_1)}{n}\right]^{\frac{1}{2}} \tag{2.46}$$

$$\frac{1}{(2\pi)^{d/2}} \tag{2.47}$$

$$\left[\frac{m(L_2-1)}{L_2^2}\right]^{\frac{K_1(K_2+1)}{2}} \left[\frac{1}{p_2(1-q_2)}\right]^{\frac{K_1}{2}} \left[\frac{L_2^2 K_2}{m^2} + \frac{(1-q_2+p_2)}{m}\right]^{\frac{K_1}{2}} \tag{2.48}$$

$$\exp\left(-\frac{nm}{2L_1^2}\left[\frac{L_1 z_1^2(1)}{np_1} + \frac{L_1 z_1^2(K_1+1)}{n(1-q_1)} + \sum_{j=1}^{K_1} (z_1(j+1)-z_1(j))^2\right]\right) \tag{2.49}$$

$$\exp\left\{-\frac{m(L_2-1)}{2L_2^2} \sum_{j=1}^{K_1} \left[\frac{L_2 z_2^2(1,j)}{mp_2} + \frac{L_2 z_2^2(K_2+1,j)}{m(1-q_2)} + \sum_{i=1}^{K_2} (z_2(i+1,j)-z_2(i,j))^2\right]\right\} \tag{2.50}$$

2.3.2. Some Useful Elementary Facts

In this subsection we will state some results and show some typical expansions that we will use in the proofs of Section 2.

Define $\Phi(x)$ as the standard normal c.d.f. valued at x. For any event E, $\overline{E}$ denotes the complment of E.

Since for any events $E(k)$ $(1 \le k \le K)$, $P\{\cap_{k=1}^K E(k)\} + P\{\cup_{k=1}^K \overline{E}(k)\} = 1$ and $P\{\cup_{k=1}^K \overline{E}(k)\} \le \sum_{k=1}^K P\{\overline{E}(k)\}$ then

$$P\{\bigcap_{k=1}^{K} E(k)\} \ge 1 - \sum_{k=1}^{K} P\{\overline{E}(k)\} . \tag{2.51}$$

In Feller [5], page 175, it is proved that

$$\text{for any} \quad x > 0 \,, \quad 1 - \Phi(x) < \frac{1}{x} \,. \tag{2.52}$$

Recalling (2.5) and expanding by Taylor's series we have

$$F_{1n}\left(T_1(j) + \frac{W_1(j)}{\sqrt{n}t_1(j)}\right) = \frac{np_1+(j-1)L_1}{n} + \frac{W_1(j)}{\sqrt{n}} + \frac{f'_{1n}(\theta_1(j))W_1^2(j)}{2nt_1^2(j)} \tag{2.53}$$

$$(1 \le j \le K_1+1)$$

where $\theta_1(j)$ is in the open interval (B_{11}, B_{12}) and therefore $|f'_{1n}(\theta_1(j))| < D_1$.

Recalling (2.6) and expanding by Taylor's series we have

$$G_{2n}\left(T_2(i,j) + \frac{W_2(i,j)}{\sqrt{m}t_2(i,j)}\,,\, j\right) = \frac{mp_2+(i-1)L_2}{m} + \frac{W_2(i,j)}{\sqrt{m}} + \frac{g'_{2n}(\theta_2(i,j),j)W_2^2(i,j)}{2mt_2^2(i,j)}$$

$$(1 \le i \le K_2+1, 1 \le j \le K_1) \tag{2.54}$$

where $\theta_2(i,j)$ is in the open interval $(B_{21}(j), B_{22}(j))$ and therefore $|g'_{2n}(\theta_2(i,j),j)| < D_2$ for all n $(1 \le i \le K_2+1, 1 \le j \le K_1)$.

Another elementary result that will be useful is the fact that for any random variables X , Y

$$\sigma(X-Y) \le \sigma(X) + \sigma(Y) \,. \tag{2.55}$$

2.4. The Asymptotic Joint Normality of the Functions of Order Statistics

In this section we will state and prove a theorem which shows that the distribution of $W(n)$ asymptotically approaches the distribution of $Z(n)$, or using the notation of Appendix A, that the sequence $\{W(n)\}$ is asymptotically equivalent to the sequence $\{Z(n)\}$. The proof of the theorem will be based on results found by Weiss [20], [21] which are explicitly stated in Theorem A.1.

We note the fact that by assumptions (5) and (10) the dimension of $Z(n)$ and $W(n)$ increases as n increases.

Theorem 2.1.

Under the assumptions listed in Section 2.2, and if for each n, $R(n)$ is any measurable region in d-dimensional euclidean space, then

$$\lim_{n\to\infty} |P\{W(n)\in R(n)\} - P\{Z(n)\in R(n)\}| = 0 .$$

Proof of the Theorem 2.1.

For the proof of the theorem, it is sufficient to prove that $\log(g(Z(n),n))/(h(Z(n),n))$ converges stochastically to zero as n increases.

Throughout this proof we assume that the p.d.f. for $Z(n)$ is $h(z(n),n)$.

Define the following events:

$$E_1(1) = \{B_{11} < T_1(1) + \frac{Z_1(1)}{\sqrt{n}t_1(1)}\}$$

$$E_1(j) = \{T_1(j-1) + \frac{Z_1(j-1)}{\sqrt{n}t_1(j-1)} < T_1(j) + \frac{Z_1(j)}{\sqrt{n}t_1(j)}\} \qquad (2 \le j \le K_1+1)$$

$$E_1(K_1+2) = \{T_1(K_1+1) + \frac{Z_1(K_1+1)}{\sqrt{n}t_1(K_1+1)} < B_{12}\}$$

$$E_2(1,j) = \{B_{21}(j) < T_2(1,j) + \frac{Z_2(1,j)}{\sqrt{m}t_2(1,j)}\} \qquad (1 \le j \le K_1)$$

$$E_2(i,j) = \{T_2(i-1,j) + \frac{Z_2(i-1,j)}{\sqrt{m}t_2(i-1,j)} < T_2(i,j) + \frac{Z_2(i,j)}{\sqrt{m}t_2(i,j)}\}$$

$$(2 \le i \le K_2+1, 1 \le j \le K_1)$$

$$E_2(K_2+1,j) = \{T_2(K_2+1,j) + \frac{Z_2(K_2+1,j)}{\sqrt{m}t_2(K_2+1,j)} < B_{22}(j)\} \qquad (1 \le j \le K_1)$$

$$E = [\bigcap_{j=1}^{K_1+2} E_1(j)] \cap [\bigcap_{i=1}^{K_2+2} \bigcap_{j=1}^{K_1} E_2(i,j)] .$$

The complement of E is the region where $g(Z(n),n)$ is zero, and in the following lemma we are going to prove that the probability of E approaches 1 as n increases.

Lemma 2.1.

$$\lim_{n\to\infty} P\{E\} = 1 .$$

Proof of Lemma 2.1.

$$P\{E\} = 1 - P\{\bar{E}\} \geq 1 - [\sum_{j=1}^{K_1+2} P\{\bar{E}_1(j)\}] - [\sum_{i=1}^{K_2+2} \sum_{j=1}^{K_1} P\{\bar{E}_2(i,j)\}]$$

therefore to prove this lemma it is sufficient to prove that the following quantities approach zero as n increases: $P\{\bar{E}_1(1)\}$, $P\{\bar{E}_1(K_1+2)\}$, $\sum_{j=2}^{K_1+1} P\{\bar{E}_1(j)\}$, $\sum_{j=1}^{K_1} P\{\bar{E}_2(1,j)\}$, $\sum_{j=1}^{K_1} P\{\bar{E}_2(K_2+2,j)\}$, $\sum_{j=1}^{K_1} \sum_{i=2}^{K_2+1} P\{\bar{E}_2(i,j)\}$.

$$P\{\bar{E}_1(1)\} = P\{T_1(1) + \frac{Z_1(1)}{\sqrt{n}t_1(1)} \leq B_{11}\} = P\{Z_1(1) \leq \sqrt{n}t_1(1)[B_{11}-T_1(1)]\}$$

and by (2.27)

$$P\{\bar{E}_1(1)\} = P\{\frac{Z_1}{\sqrt{p_1(1+\delta_1)}} \leq \frac{\sqrt{n}t_1(1)[B_{11}-T_1(1)]}{\sqrt{p_1(1+\delta_1)}}\}$$

and by assumption (30) this approaches zero as n increases.

$$P\{\bar{E}_1(K_1+2)\} = P\{B_{12} \leq T_1(K_1+1) + \frac{Z_1(K_1+1)}{\sqrt{n}t_1(K_1+1)}\}$$

$$= P\{\frac{\sqrt{n}t_1(K_1+1)[B_{12}-T_1(K_1+1)]}{\sqrt{(1-q_1)(1+\delta_1')}} \leq \frac{Z_1(K_1+1)}{\sqrt{(1-q_1)(1+\delta_1')}}\}$$

and by (2.28) the right side is standard normal and by assumption (31) $P\{\bar{E}_1(K_1+2)\}$ approaches zero as n increases.

$$P\{\bar{E}_1(j+1)\} = P\{\frac{Z_1(j+1)}{t_1(j+1)} - \frac{Z_1(j)}{t_1(j)} \leq \sqrt{n}[T_1(j)-T_1(j+1)]\} \qquad (1 \leq j \leq K_1)$$

and we have $\sqrt{n}[T_1(j)-T_1(j+1)] = (-L_1)/(\sqrt{n}f_{1n}(W_1'(j))$ where $W_1'(j)$ is a real number in the open interval $(T_1(1),T_1(K_1+1))$ and therefore by (2.16)

$$P\{\bar{E}_1(j+1)\} \leq P\{\frac{Z_1(j+1)}{t_1(j+1)} - \frac{Z_1(j)}{t_1(j)} \leq \frac{-L_1}{\sqrt{n}B_1}\} \qquad (1 \leq j \leq K_1)$$

and since

$$\frac{Z_1(j+1)}{t_1(j+1)} - \frac{Z_1(j)}{t_1(j)} = \frac{Z_1(j+1)}{t_1(j) + \frac{L_1 f_{1n}'(W_1''(j))}{n\, f_{1n}(W_1''(j))}} - \frac{Z_1(j)}{t_1(j)}$$

$$= \frac{Z_1(j+1)-Z_1(j)}{t_1(j)} - \frac{Z_1(j+1)\frac{L_1 f_{1n}'(W_1''(j))}{n\, f_{1n}(W_1''(j))t_1(j)}}{1 + \frac{L_1 f_{1n}'(W_1''(j))}{n\, f_{1n}(W_1''(j))t_1(j)}}$$

therefore by (2.55), (2.33), (2.29) and (2.15) we have

$$\sigma(\frac{Z_1(j+1)}{t_1(j+1)} - \frac{Z_1(j)}{t_1(j)}) \leq \frac{1}{b_1}\sqrt{\frac{L_1}{n}(1+\delta_1'')} + \left(\frac{\frac{L_1 D_1}{nb_1^2}}{1 + \frac{L_1 D_1}{nb_1^2}}\right)\bar{\sigma} = \beta_1$$

say, and by assumptions (12) and (16) β_1 approaches zero as n increases.

Therefore,

$$P\{\bar{E}_1(j+1)\} \leq \Phi(\frac{-L_1}{\sqrt{n}B_1\beta_1}) = 1 - \Phi(\frac{L_1}{\sqrt{n}B_1\beta_1}) < \frac{\sqrt{n}B_1\beta_1}{L_1}$$

by (2.52). It follows that

$$\sum_{j=2}^{K_1+1} P\{\bar{E}_1(j)\} < \frac{\sqrt{n}B_1K_1\beta_1}{L_1}$$

and by assumptions (17) and (18) this converges to zero as n increases.

$$\sum_{j=1}^{K_1} P\{\bar{E}_2(1,j)\} = \sum_{j=1}^{K_1} \Phi(\frac{\sqrt{m}t_2(1,j)}{\sqrt{P_2(1+\delta_2)}}[B_{21}(j)-T_2(1,j)])$$

by (2.30) therefore,

$$\sum_{j=1}^{K_1} P\{\bar{E}_2(1,j)\} = \sum_{j=1}^{K_1} [1 - \Phi(\frac{\sqrt{m}t_2(1,j)}{\sqrt{P_2(1+\delta_2)}}|B_{21}(j)-T_2(1,j)|)]$$

$$\leq \max_{1\leq j\leq K_1} \frac{K_1\sqrt{P_2}\sqrt{1+\delta_2}}{\sqrt{m}b_2|B_{21}(j)-T(1,j)|}$$

and by assumption (32) this last expression converges to zero as n increases. Using (2.31) and (2.52)

$$\sum_{j=1}^{K_1} P\{\bar{E}_2(K_2+2,j)\} = \sum_{j=1}^{K_1} [1 - P\{E_2(K_2+2,j)\}] =$$

$$= \sum_{j=1}^{K_1} [1 - \Phi(-\frac{\sqrt{m}t_2(K_2+1,j)}{\sqrt{1-q_2}\sqrt{1+\delta_2'}} [B_{22}(j)-T_2(K_2+1,j)])]$$

$$< \sum_{j=1}^{K_1} \frac{\sqrt{1-q_2}\sqrt{1+\delta_2'}}{\sqrt{m}t_2(K_2+1,j)[B_{22}(j)-T_2(K_2+1,j)]}$$

$$\leq \max_{1\leq j\leq K_1} \frac{\sqrt{1-q_2}\sqrt{1+\delta_2'}\ K_1}{\sqrt{m}b_2[B_{22}(j)-T_2(K_2+1,j)]}$$

and by assumption (33) this converges to zero as n increases.

$$P\{\bar{E}_2(i+1,j)\} = P\{\frac{Z_2(i+1,j)}{t_2(i+1,j)} - \frac{Z_2(i,j)}{t_2(i,j)} \leq \sqrt{m}[T_2(i,j)-T_2(i+1,j)]\}$$

$$(1\leq i\leq K_2, 1\leq j\leq K_1)$$

and we have $\sqrt{m}[T_2(i,j)-T_2(i+1,j)] = (-L_2)/(\sqrt{m}g_{2n}(W_2'(i,j),j))$ where $W_2'(i,j)$ is in the open interval $(T_2(1,j),T_2(K_2+1,j))$ and therefore by (2.18)

$$P\{\bar{E}_2(i+1,j)\} = P\{\frac{Z_2(i+1,j)}{t_2(i+1,j)} - \frac{Z_2(i,j)}{t_2(i,j)} \leq \frac{-L_2}{\sqrt{m}B_2}\} \qquad (1\leq i\leq K_2, 1\leq j\leq K_1)$$

and since

$$\frac{Z_2(i+1,j)}{t_2(i+1,j)} - \frac{Z(i,j)}{t_2(i,j)} = \frac{Z_2(i+1,j)}{t_2(i+1,j) + \dfrac{L_2 g'_{2n}(W''_2(i,j),j)}{m\, g_{2n}(W''_2(i,j),j)}} - \frac{Z_2(i,j)}{t_2(i,j)}$$

$$= \frac{Z_2(i+1,j)-Z_2(i,j)}{t_2(i,j)} - \frac{Z_2(i+1,j)\,\dfrac{L_2 g'_{2n}(W''_2(i,j),j)}{m g_{2n}(W''_2(i,j),j)t_2(i,j)}}{1 + \dfrac{L_2 g'_{2n}(W''_2(i,j),j)}{m\, g_{2n}(W''_2(i,j))t_2(i,j)}}$$

then by (2.55), (2.33), (2.32) and (2.17)

$$\sigma\left(\frac{Z_2(i+1,j)}{t_2(i+1,j)} - \frac{Z_2(i,j)}{t_2(i,j)}\right) \leq \frac{1}{b_2}\sqrt{\frac{L_2}{m}(1+\delta''_2)} + \left(\frac{\dfrac{L_2 D_2}{m b_2^2}}{1 + \dfrac{L_2 D_2}{m b_2^2}}\right) \bar{\sigma} = \beta_2$$

say and be assumptions (13) and (23) β_2 approaches zero as n increases. Therefore

$$P\{\bar{E}_2(i+1,j)\} \leq \Phi\left(\frac{-L_2}{\sqrt{m}B_2\beta_2}\right) = 1 - \Phi\left(\frac{L_2}{\sqrt{m}B_2\beta_2}\right) < \frac{\sqrt{m}B_2\beta_2}{L_2}$$

by (2.52) $(1\leq i\leq K_2, 1\leq j\leq K_1)$, therefore

$$\sum_{j=1}^{K_1}\sum_{i=2}^{K_2+1} P\{\bar{E}(i,j)\} < \frac{K_1(K_2+1)\sqrt{m}B_2\beta_2}{L_2}$$

and by assumptions (24) and (25) this converges to zero as n increases, and this completes the proof of Lemma 2.1.

Lemma 2.2.

With probability approaching one as n increases we may write $\log(g(Z(n),n))/(h(Z(n),n))$ as the sum of the following sixteen expressions

$$(np_1-1)\log[F_{1n}(T_1(1)+\frac{Z_1(1)}{\sqrt{n}t_1(1)})] \tag{2.56}$$

$$(mp_2-1)\sum_{j=1}^{K_1}\log[G_{2n}(T_2(1,j)+\frac{Z_2(1,j)}{\sqrt{m}t_2(1,j)},j)] \tag{2.57}$$

$$n(1-q_1)\log[1-F_{1n}(T_1(K_1+1)+\frac{Z_1(K_1+1)}{\sqrt{n}t_1(K_1+1)})] \tag{2.58}$$

$$m(1-q_2)\sum_{j=1}^{K_1}\log[1-G_{2n}(T_2(K_2+1,j)+\frac{Z_2(K_2+1,j)}{\sqrt{m}t_2(K_2+1,j)},j)] \tag{2.59}$$

$$m\sum_{j=1}^{K_1}\log[F_{1n}(T_1(j+1)+\frac{Z_1(j+1)}{\sqrt{n}t_1(j+1)})-F_{1n}(T_1(j)+\frac{Z_1(j)}{\sqrt{n}t_1(j)})] \tag{2.60}$$

$$(L_2-1)\sum_{j=1}^{K_1}\sum_{i=1}^{K_2}\log[G_{2n}(T_2(i+1,j)+\frac{Z_2(i+1,j)}{\sqrt{m}t_2(i+1,j)},j)$$

$$-G_{2n}(T_2(i,j)+\frac{Z_2(i,j)}{\sqrt{m}t_2(i,j)},j)] \tag{2.61}$$

$$\frac{nm}{2L_1^2}[\frac{L_1Z_1^2(1)}{np_1}+\frac{L_1Z_1^2(K_1+1)}{n(1-q_1)}+\sum_{j=1}^{K_1}(Z_1(j+1)-Z_1(j))^2] \tag{2.62}$$

$$\frac{m(L_2-1)}{2L_2^2}\sum_{j=1}^{K_1}\left[\frac{L_2Z_2^2(1,j)}{mp_2}+\frac{L_2Z_2^2(K_2+1,j)}{m(1-q_2)}+\sum_{i=1}^{K_2}(Z_2(i+1,j)-Z_2(i,j))^2\right] \tag{2.63}$$

$$-\sum_{j=1}^{K_1+1}\log t_1(j) \tag{2.64}$$

$$-\sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1}\log t_2(i,j) \tag{2.65}$$

$$\sum_{j=1}^{K_1+1}\log f_{1n}(T_1(j)+\frac{Z_1(j)}{\sqrt{n}t_1(j)}) \tag{2.66}$$

$$\sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1}\log g_{2n}(T_2(i,j)+\frac{Z_2(i,j)}{\sqrt{m}t_2(i,j)}) \tag{2.67}$$

$$\log\left[\frac{n!}{(np_1-1)!(n-nq_1)![m!]^{K_1}n^{(K_1+1)/2}}\right] \tag{2.68}$$

$$K_1\log\left[\frac{m!}{(mp_2-1)!(m-mq_2)![(L_2-1)!]^{K_2}m^{(K_2+1)/2}}\right] \tag{2.69}$$

$$\frac{(K_1+1)}{2}\log 2\pi-\log\left[\left(\frac{nm}{L_1^2}\right)^{\frac{K_1+1}{2}}(P_1(1-q_1))^{-\frac{1}{2}}\left(\frac{L_1^2K_1}{n^2}+\frac{L_1(1-q_1+p_1)}{n}\right)^{\frac{1}{2}}\right] \tag{2.70}$$

$$\frac{K_1(K_2+1)}{2}\log 2\pi-\log\left[\left(\frac{m(L_2-1)}{L_2^2}\right)^{(K_1(K_2+1))/2}(P_2(1-q_2))^{(-K_1)/2}\cdot\left(\frac{L_2^2K_2}{m^2}+\frac{L_2(1-q_2+p_2)}{m}\right)^{(K_1)/2}\right]. \tag{2.71}$$

Proof of Lemma 2.2.

Applying Lemma 2.1 and the formula for $g(w(n),n)$ and $h(z(n),n)$ it follows immediately that Lemma 2.2 holds.

Lemma 2.3.

The expression (2.56) can be written as

$$(np_1-1)\log p_1 + np_1\left[\frac{Z_1(1)}{\sqrt{n}p_1} + \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{2np_1t_1^2(1)} - \frac{Z_1^2(1)m}{2np_1^2L_1}\right] + \Delta(1,n) \qquad (2.56)^*$$

where $\Delta(1,n)$ converges stochastically to zero as n increases.

Proof of Lemma 2.3.

Using (2.53), (2.56) can be written as

$$(np_1-1)\log\left[p_1 + \frac{Z_1(1)}{\sqrt{n}} + \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{2nt_1^2(1)}\right]$$

where $\theta_1(1)$ is a real number in the open interval (B_{11},B_{12}) , and therefore $|f'_{1n}(\theta_1)| < D_1$ by assumption (12). Then (2.56) can be written as

$$(np_1-1)\log p_1 + (np_1-1)\log\left[1 + \frac{Z_1(1)}{\sqrt{n}\ p_1} + \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{2np_1t_1^2(1)}\right]$$

and since by (2.27) $(Z_1(1))/(\sqrt{p_1(1+\delta_1)})$ is standard normal and by assumption (16) $\lim_{n\to\infty} (L_1D_1)/(nb_1^2) = 0$ it follows that

$$\frac{Z_1(1)}{\sqrt{n}\,p_1} + \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{2np_1t_1^2(1)}$$

converges stochastically to zero as n increases. Then, using a Taylor expansion, (2.56) can be written as

$$(np_1-1)\log p_1 + (np_1-1)\left[\frac{Z_1(1)}{\sqrt{n}\,p_1} + \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{2np_1t_1^2(1)}\right.$$

$$\left. - \frac{1}{2}(1+\gamma_1)\left[\frac{Z_1(1)}{\sqrt{n}\,p_1} + \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{2np_1t_1^2(1)}\right]^2\right]$$

where γ_1 converges stochastically to zero as n increases. Recalling assumptions (3) and (19) the proof of Lemma 2.3 is completed.

Lemma 2.4.

The expression (2.57) can be written as

$$(mp_2-1)K_1 \log p_2 + mp_2 \sum_{j=1}^{K_1} \left[\frac{Z_2(1,j)}{\sqrt{m}\,p_2} + \frac{g'_{2n}(\theta_2(1,j),j)Z_2^2(1,j)}{2mp_2t_2^2(1,j)}\right.$$

$$\left. - \frac{Z_2^2(1,j)(L_2-1)}{2mp_2^2L_2}\right] + \Delta(2,n) \qquad (2.57)^*$$

where $\Delta(2,n)$ converges stochastically to zero as n increases.

Proof of Lemma 2.4.

Using (2.54) the expression (2.57) can be written as

$$(mp_2-1)\sum_{j=1}^{K_1}\log[p_2+\frac{Z_2(1,j)}{\sqrt{m}}+\frac{g'_{2n}(\theta_2(1,j),j)Z_2^2(1,j)}{2mt_2^2(1,j)}]$$

where $\theta_2(1,j)$ is in the open interval $(B_{21}(j),B_{22}(j))$ $(1\leq j\leq K_1)$, and therefore $|g'_{2n}(\theta_2(1,j),j)|<D_2$ $(1\leq j\leq K_1)$. Then (2.57) can be written as

$$(mp_2-1)K_1\log p_2+(mp_2-1)\sum_{j=1}^{K_1}[1+\frac{Z_2(1,j)}{\sqrt{m}\,p_2}+\frac{g'_{2n}(\theta_2(1,j),j)Z_2^2(1,j)}{2mp_2t_2^2(1,j)}],$$

and since by (2.30), $(Z_2(1,j))/(\sqrt{p_2(1+\delta_2)})$ is standard normal and $\lim_{n\to\infty}(L_2D_2)/(mb_2^2)=0$ by assumption (23), it follows that

$$\frac{Z_2(1,j)}{\sqrt{m}\,p_2}+\frac{g'_{2n}(\theta_2(1,j),j)Z_2^2(1,j)}{2mp_2t_2^2(1,j)}$$

converges stochastically to zero as n increases. Then using a Taylor expansion we can write (2.57) as

$$(mp_2-1)K_1\log p_2+(mp_2-1)\sum_{j=1}^{K_1}[\frac{Z_2(1,j)}{\sqrt{m}\,p_2}+\frac{g'_{2n}(\theta_2(1,j),j)Z_2^2(1,j)}{2mp_2t_2^2(1,j)}$$

$$-\frac{1}{2}(1+\gamma_2(j))[\frac{Z_2(1,j)}{\sqrt{m}\,p_2}+\frac{g'_{2n}(\theta_2(1,j),j)Z_2(1,j)}{2mp_2t_2^2(1,j)}]^2]$$

where $\gamma_2(j)$ $(1\leq j\leq K_1)$ converges stochastically to zero as n increases, and recalling assumptions (8) and (26) the proof of Lemma 2.4 is completed.

Lemma 2.5.

The expression (2.58) can be written as

$$n(1-q_1)\log(1-q_1) + n(1-q_1)\left[\frac{-Z_1(K_1+1)}{\sqrt{n}(1-q_1)} - \frac{mZ_1^2(K_1+1)}{2n(1-q_1)^2 L_1} - \frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_1+1)}{2n(1-q_1)^2 t_1^2(K_1+1)}\right] + \Delta(3,n) \qquad (2.58)^*$$

where $\Delta(3,n)$ converges stochastically to zero as n increases.

Proof of Lemma 2.5.

Using (2.53) the expression (2.58) can be written as

$$n(1-q_1)\log\left[1 - q_1 - \frac{Z_1(K_1+1)}{\sqrt{n}} - \frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_1+1)}{2nt_1^2(K_1+1)}\right]$$

where $\theta_1(K_1+1)$ is a real number in the open interval (B_{11},B_{12}), and therefore $|f'_{1n}(\theta_1(K_1+1))| < D_1$. Then (2.58) can be written as

$$n(1-q_1)\log(1-q_1) + n(1-q_1)\log\left[1 - \frac{Z_1(K_1+1)}{\sqrt{n}(1-q_1)} - \frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_1+1)}{2n(1-q_1)t_1^2(K_1+1)}\right],$$

and since by (2.28), $(Z_1(K_1+1))/(\sqrt{(1-q_1)(1+\delta_1')})$ is standard normal

and by assumption (16), $\lim_{n\to\infty} (L_1D_1)/(nb_1^2) = 0$ it follows that

$$\frac{Z_1(K_1+1)}{\sqrt{n}(1-q_1)} + \frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_1+1)}{2n(1-q_1)t_1^2(K_1+1)}$$

converges stochastically to zero as n increases. By a Taylor expansion we have that (2.58) can be written as

$$n(1-q_1)\log(1-q_1) + n(1-q_1)\left[\frac{-Z_1(K_1+1)}{\sqrt{n}(1-q_1)} - \frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_1+1)}{2n(1-q_1)t_1^2(K_1+1)}\right.$$

$$\left. - \frac{1}{2}(1+\gamma_{K_1+1})\left[\frac{Z_1(K_1+1)}{\sqrt{n}(1-q_1)} + \frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_i+1)}{2n(1-q_1)t_1^2(K_1+1)}\right]^2\right]$$

where γ_{K_1+1} converges stochastically to zero as n increases. Recalling assumptions (4) and (20) the proof of Lemma 2.5 is completed.

<u>Lemma 2.6</u>.

The expression (2.59) can be written as

$$mK_1(1-q_2)\log(1-q_2) + m(1-q_2)\sum_{j=1}^{K_1}\left[\frac{-Z_2(K_2+1,j)}{\sqrt{m}t_2(K_2+1,j)} - \frac{Z_2^2(K_2+1,j)(L_2-1)}{2m(1-q_2)L_2} - \frac{f'_{2n}(\theta_2(K_2+1,j),j)Z_2^2(K_2+1,j)}{2m(1-q_2)t_2^2(K_2+1,j)}\right]$$

$$+ \Delta(4,n) \qquad (2.59)^*$$

where $\Delta(4,n)$ converges stochastically to zero as n increases.

Proof of Lemma 2.6.

Using (2.54) and recalling the fact that $K_2L_2 = m(q_2-p_2)$, (2.59) can be written as

$$m(1-q_2)\sum_{j=1}^{K_1}\log[1 - q_2 - \frac{Z_2(K_2+1,j)}{\sqrt{m}} - \frac{g'_{2n}(\theta_2(K_2+1,j),j)Z_2^2(K_2+1,j)}{2mt_2^2(K_2+1,j)}]$$

where $\theta_2(K_2+1,j)$ is in the open interval $(B_{21}(j),B_{22}(j))$ $(1\leq j\leq K_1)$ and therefore $|g'_{2n}(\theta_2(K_2+1,j),j)| < D_2$ $(1\leq j\leq K_1)$. (2.59) can be written as

$$mK_1(1-q_2)\log(1-q_2)$$

$$+ m(1-q_2)\sum_{j=1}^{K_1}\log[1 - \frac{Z_2(K_2+1,j)}{\sqrt{m}(1-q_2)} - \frac{g'_{2n}(\theta_2(K_2+1,j),j)Z_2^2(K_2+1,j)}{2m(1-q_2)t_2^2(K_2+1,j)}] .$$

Since by (2.31), $(Z_2(K_2+1,j))/(\sqrt{(1-q_2)(1+\delta_2')})$ is standard normal and by assumption (23), $\lim_{n\to\infty} (L_2D_2)/(mb_2^2) = 0$ it follows that

$$\frac{Z_2(K_2+1,j)}{\sqrt{m}(1-q_2)} + \frac{g'_{2n}(\theta_2(K_2+1,j)Z_2^2(K_2+1,j)}{2m(1-q_2)t_2^2(K_2+1,j)}$$

converges stochastically to zero as n increases. Therefore, using a Taylor expansion (2.59) may be written as

$$m(1-q_2)K_1 \log(1-q_2)$$

$$+ m(1-q_2) \sum_{j=1}^{K_1} \left[\frac{-Z_2(K_2+1,j)}{\sqrt{m}(1-q_2)} - \frac{g'_{2n}(\theta_2(K_2+1,j),j)Z_2^2(K_2+1,j)}{2m(1-q_2)t_2^2(K_2+1,j)} \right.$$

$$- \frac{1}{2}(1+\gamma_2(K_2+1,j))[\frac{Z_2(K_2+1,j)}{\sqrt{m}(1-q_2)}$$

$$\left. + \frac{g'_{2n}(\theta_2(K_2+1,j),j)Z_2^2(K_2+1,j)}{2m(1-q_2)t_2^2(K_2+1,j)}]^2 \right]$$

where $\gamma_2(K_2+1,j)$ converges stochastically to zero as n increases. Recalling assumptions (9) and (27) the proof of Lemma 2.6 is completed.

<u>Lemma 2.7.</u>

The expression (2.60) can be written as

$$(L_1-1)K_1 \log \frac{L_1}{n} + \sqrt{n}(Z_1(K_1+1)-Z_1(1)) - \frac{n(L_1-1)}{2L_1^2} \sum_{j=1}^{K_1} [Z_1(j+1)-Z_1(j)]^2$$

$$(2.60)^*$$

$$+ \frac{1}{2}[\frac{f'_{1n}(\theta_1(K_1+1))Z_1^2(K_1+1)}{t_1^2(K_1+1)} - \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{t_1^2(1)}] + \Delta(5,n)$$

where $\Delta(5,n)$ converges stochastically to zero as n increases.

<u>Proof of Lemma 2.7.</u>

Using (2.53) we can write (2.60) as

$$= \sum_{j=1}^{K_1} \log\left[\frac{L_1}{n} + \frac{Z_1(j+1)-Z_1(j)}{\sqrt{n}} + \frac{Q_1(j)}{2n}\right]$$

where $Q_1(j)$ is defined as

$$\frac{f'_{1n}(\theta_1(j+1))Z_1^2(j+1)}{t_1^2(j+1)} - \frac{f'_{1n}(\theta_1(j+1))Z_1^2(j)}{t_1^2(j)} \qquad (1 \le j \le K_1) \ .$$

Then we can write (2.60) as

$$mK_1 \log\frac{L_1}{n} + m\sum_{j=1}^{K_1} \log\left[1 + \frac{\sqrt{n}(Z_1(j+1)-Z_1(j))}{L_1} + \frac{Q_1(j)}{2L_1}\right] \ .$$

we are going to show that

$$\max_{1 \le j \le K_1} \left|\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1} + \frac{Q_1(j)}{2L_1}\right|$$

converges stochastically to zero as n increases. Since

$$\max_{1 \le j \le K_1} \left|\frac{\sqrt{n}(Z_1(j+1)-Z_1(j))}{L_1} + \frac{Q_1(j)}{2L_1}\right|$$

$$\le \max_{1 \le j \le K_1} \left|\frac{\sqrt{n}(Z_1(j+1)-Z_1(j))}{L_1}\right| + \max_{1 \le j \le K_1} \left|\frac{Q_1(j)}{2L_1}\right|$$

it is sufficient to prove that both terms on the right side of the inequality converge stochastically to zero. Fix a positive number ε . Using (2.51) and (2.52)

$$P\{\max_{1\le j\le K_1}\left|\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1}\right| \le \varepsilon\} \ge 1 - \sum_{j=1}^{K_1} P\{\left|\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1}\right| > \varepsilon\}$$

$$= 1 - 2\sum_{j=1}^{K_1} P\{\frac{Z_1(j+1)-Z_1(j)}{\sqrt{\frac{L_1}{n}(1+\delta_1'')}} > \frac{\varepsilon\sqrt{L_1}}{\sqrt{1+\delta_1''}}\}$$

$$= 1 - 2\sum_{j=1}^{K_1} [1 - \Phi(\frac{\varepsilon\sqrt{L_1}}{\sqrt{1+\delta_1''}})]$$

$$> 1 - \frac{2K_1\sqrt{1+\delta_1''}}{\varepsilon\sqrt{L_1}} .$$

By assumption (21), $\lim_{n\to\infty} (K_1\sqrt{D_1})/(b_1\sqrt{L_1}) = 0$ and since by assumption (12), $D_1 > 1$ and by assumption (14), $b_1 \le 1$ for all sufficiently large n it follows that $\lim_{n\to\infty} K_1/\sqrt{L_1} = 0$ and therefore the proof that

$$\max_{1\le j\le K_1}\left|\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1}\right|$$

converges stochastically to zero, as n increases, is completed. We are now going to show that

$$\max_{1\le j\le K_1}\left|\frac{Q_1(j)}{L_1}\right|$$

converges stochastically to zero as n increases. We note the fact that

$$\max_{1\le j\le K_1} \left|\frac{Q_1(j)}{L_1}\right| \le \frac{2D_1 \max_{1\le j\le K_1+1} Z_1^2(j)}{L_1 b_1^2} .$$

$$P\left\{\frac{2D_1 \max_{1\le j\le K_1+1} Z_1^2(j)}{L_1 b_1^2} \le \varepsilon\right\} = P\left\{\max_{1\le j\le K_1+1} Z^2{}_1(j) \le \frac{\varepsilon L_1 b_1^2}{2D_1}\right\}$$

$$> 1 - \sum_{j=1}^{K_1+1} P\{Z_1^2(j) > \frac{\varepsilon L_1 b_1^2}{2D_1}\}$$

$$= 1 - 2 \sum_{j=1}^{K_1+1} P\{\frac{Z_1(j)}{\sigma(Z_1(j))} > \frac{\sqrt{\varepsilon}\sqrt{L_1}b_1}{\sqrt{2D_1}\sigma(Z_1(j))}\}$$

$$> 1 - 2 \sum_{j=1}^{K_1+1} [1 - \Phi(\frac{\sqrt{\varepsilon}\sqrt{L_1}b_1}{\sqrt{2D_1}\sigma(Z_1(j))}$$

$$> 1 - \frac{2(K_1+1)\sqrt{2D_1}\ \bar{\sigma}}{\sqrt{\varepsilon}\sqrt{L_1}\ b_1}$$

and by assumption (21), $\lim_{n\to\infty} (K_1\sqrt{D_1})/(\sqrt{L_1}b_1) = 0$ and therefore the proof that

$$\max_{1\le j\le K_1} \left|\frac{Q_1(j)}{L_1}\right|$$

converges stochastically to zero is completed. Therefore we may write (2.60) using a Taylor expansion as

$$= mK_1 \log \frac{L_1}{n} + m \sum_{j=1}^{K_1} \left[\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1} + \frac{Q_1(j)}{2L_1} - \frac{1}{2}\left[\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1} + \frac{Q_1(j)}{2L_1}\right]^2 + \frac{1}{3(1+\gamma_j)^3}\left[\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1} + \frac{Q_1(j)}{2L_1}\right]^3 \right]$$

where γ_j $(1 \leq j \leq K_1)$ converges stochastically to zero as n increases. The last expression is (2.60)* where

$$\Delta(5,n) = -\frac{\sqrt{n}[Z_1(K_1+1)-Z_1(1)]}{L_1}$$

$$m \sum_{j=1}^{K_1} \left\{ -\frac{1}{8L_1^2} Q_1^2(j) - \frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]Q_1(j)}{2L_1^2} + \frac{1}{3(1+\gamma_j)^3}\left[\frac{\sqrt{n}[Z_1(j+1)-Z_1(j)]}{L_1} + \frac{Q_1(j)}{2L_1}\right]^3 \right\}$$

$$-\frac{1}{2L_1}\left[\frac{f'_{1n}(\theta_{K_1+1})Z_1^2(K_1+1)}{t^2(K_1+1)} - \frac{f'_{1n}(\theta_1(1))Z_1^2(1)}{t_1^2(1)}\right].$$

To show that $\Delta(5,n)$ converges stochastically to zero as n increases it is sufficient to show that

$$R_1(r,s) = m \sum_{j=1}^{K_1} \left[\left(\frac{\sqrt{n}(Z_1(j+1)-Z_1(j))}{L_1}\right)^r \left(\frac{Q_1(j)}{L_1}\right)^s \right]$$

converges stochastically to zero for the following pairs: (r,s): $(1,1)$, $(0,2),(0,3),(1,2),(2,1),(3,0)$. We have

$$|R_1(r,s)| \le \frac{mD_1^s}{L_1^s b_1^{2s}(\sqrt{L_1})^r} \sum_{j=1}^{K_1} \left| \frac{Z_1(j+1)-Z_1(j)}{\sqrt{\frac{L_1}{n}}} \right|^r [Z_1^2(j+1)+Z_1^2(j)]^s$$

and therefore

$$E(|R_1(r,s)|) \le \frac{K_1 m D_1^s}{L_1^{s+(r/2)} b_1^{2s}} J_1(r,s) \le \frac{nD_1^s}{L_1^{s+(r/s)} b_1^{2s}} J_1(r,s)$$

where $J_1(r,s)$ is a finite number depending only on (r,s).

By (2.22) the right side of the inequality approaches zero as n increases for all the pairs (r,s) we are interested in. Therefore $|R_1(r,s)|$ converges stochastically to zero as n increases and the proof of Lemma 2.7 is completed.

<u>Lemma 2.8.</u>

The expression (2.61) can be written as

$$(L_2-1)K_1K_2 \log \frac{L_2}{m} + \sqrt{m} \sum_{j=1}^{K_1} [Z_2(K_2+1,j)-Z_2(1,j)]$$

$$- \frac{m(L_2-1)}{2L_2^2} \sum_{j=1}^{K_1} \sum_{i=1}^{K_2} [Z_2(i+1,j)-Z_2(i,j)]^2$$

$$+ \frac{1}{2} \sum_{j=1}^{K_1} \left[\frac{g'_{2n}(\theta_2(K_2+1,j)Z_2^2(K_2+1,j)}{t_2^2(K_2+1,j)} - \right.$$

$$- \frac{g'_{2n}(\theta_2(i,j),j)Z_2^2(i,j)}{t_2^2(i,j)}\Bigg] + \Delta(6,n) \qquad (2.61)^*$$

where $\Delta(6,n)$ converges stochastically to zero as n increases.

Proof of Lemma 2.8.

Using (2.54) we can write (2.61) as

$$(L_2-1)\sum_{j=1}^{K_2}\sum_{i=1}^{K_2}\log\left[\frac{L_2}{m}+\frac{Z_2(i+1,j)-Z_2(i,j)}{\sqrt{m}}+\frac{Q_2(i,j)}{2m}\right]$$

where $Q_2(i,j)$ is defined as

$$\frac{g'_{2n}(\theta_2(i+1,j),j)Z_2^2(i+1,j)}{t_2^2(i+1,j)} - \frac{g'_{2n}(\theta_2(i,j),j)Z_2^2(i,j)}{t_2^2(i,j)} \qquad (1\leq i\leq K_2,\, 1\leq j\leq K_1)$$

and we can rewrite (2.61) as

$$(L_2-1)K_1K_2\log\frac{L_2}{m} + (L_2-1)\sum_{j=1}^{K_1}\sum_{i=1}^{K_2}\left[1+\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2}+\frac{Q_2(i,j)}{2L_2}\right].$$

We are going to show that

$$\max_{1\leq i\leq K_2,\,1\leq j\leq K_1}\left|\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)}{L_2}+\frac{Q_2(i,j)}{2L_2}\right|$$

converges stochastically to zero as n increases. It is sufficient to show that both

$$\max_{1\le j\le K_1,\,1\le i\le K_2}\left|\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2}\right|$$

and

$$\max_{1\le j\le K_1,\,1\le i\le K_2}\left|\frac{Q_2(i,j)}{2L_2}\right|$$

converge stochastically to zero as n increases. Fix a positive number ε . Using (2.51) and (2.52)

$$P\{\max_{1\le i\le K_2,\,1\le j\le K_1}\left|\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2}\right| \le \varepsilon\}$$

$$\ge 1 - \sum_{j=1}^{K_1}\sum_{i=1}^{K_2} P\{\left|\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2}\right| > \varepsilon\}$$

$$= 1 - 2\sum_{j=1}^{K_1}\sum_{i=1}^{K_2} P\{\frac{Z_2(i+1,j)-Z_2(i,j)}{\sqrt{\frac{L_2}{m}}\sqrt{1+\delta_2''}} > \frac{\varepsilon\sqrt{L_2}}{\sqrt{1+\delta_2''}}\}$$

$$= 1 - 2\sum_{j=1}^{K_1}\sum_{i=1}^{K_2} [1 - \Phi(\frac{\varepsilon\sqrt{L_2}}{\sqrt{1+\delta_2''}})]$$

$$> 1 - \frac{2K_1K_2\sqrt{1+\delta_2''}}{\varepsilon\sqrt{L_2}}$$

and since by assumption (28), $\lim_{n\to\infty} (K_1K_2\sqrt{D_2})/(\sqrt{L_2}b_2) = 0$, and by assumption (13), $D_2 > 1$ and by assumption (15), $b_2 \le 1$ for all

sufficiently large n, it follows that $\lim_{n\to\infty} (K_1K_2)/(\sqrt{L_2}) = 0$, and since the inequality holds for all positive ε, the proof that

$$\max_{1\le i\le K_2,\,1\le j\le K_1} \left| \frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2} \right|$$

converges stochastically to zero is completed.

We are now going to show that

$$\max_{1\le i\le K_2,\,1\le j\le K_1} \left| \frac{Q_2(i,j)}{2L_2} \right|$$

converges stochastically to zero as n increases. Recalling the definition of $Q_2(i,j)$ we note the fact that

$$\max_{1\le i\le K_2,\,1\le j\le K_1} \left| \frac{Q_2(i,j)}{2L_2} \right| \le \frac{D_2 \max_{1\le i\le K_2+1,\,1\le j\le K_1} Z_2^2(i,j)}{L_2 b_2^2} .$$

For any fixed positive number ε, using (2.51) and (2.52)

$$P\left\{ \frac{D_2 \max\limits_{1\le i\le K_2+1,\,1\le j\le K_1} Z_2^2(i,j)}{L_2\, b_2^2} \le \varepsilon \right\} \ge 1 - \sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} P\{Z_2^2(i,j) > \frac{\varepsilon L_2 b_2^2}{D_2}\}$$

$$= 1 - 2 \sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} P\{\frac{Z_2(i,j)}{\sigma(Z_2(i,j))} > \frac{\sqrt{\varepsilon}\sqrt{L_2}\, b_2}{\sqrt{D_2}\sigma(Z_2(i,j))}\}$$

$$\ge 1 - 2 \sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} P\{\frac{Z_2(i,j)}{\sigma(Z_2(i,j))} > \frac{\sqrt{\varepsilon}\sqrt{L_2}\, b_2}{\sqrt{D_2}\, \bar{\sigma}}\} =$$

$$= 1 - 2K_1(K_2+1)\left[1 - \Phi\left(\frac{\sqrt{\epsilon}\sqrt{L_2}\, b_2}{\sqrt{D_2}\,\bar{\sigma}}\right)\right]$$

$$> 1 - \frac{2K_1(K_2+1)\sqrt{D_2}\,\bar{\sigma}}{\sqrt{\epsilon}\sqrt{L_2}\, b_2}$$

and by assumption (28) the last expression approaches one as n increases, and since this holds for any fixed positive number ϵ, the proof that

$$\max_{1\le j\le K_1,\,1\le i\le K_2} \left|\frac{Q_2(i,j)}{2L_2}\right|$$

converges stochastically to zero is completed.

Thus we can write (2.61) as

$$(L_2-1)K_2K_1 \log \frac{L_2}{m} + (L_2-1)\sum_{j=1}^{K_1}\sum_{i=1}^{K_2}\left[\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2} + \frac{Q_2(i,j)}{2L_2} - \frac{1}{2}\left[\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2} + \frac{Q_2(i,j)}{2L_2}\right]^2 + \frac{1}{3(1+\gamma_2(i,j))^3}\cdot\left[\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{L_2} + \frac{Q_2(i,j)}{2L_2}\right]^3\right]$$

where $\max_{1\le i\le K_2,\,1\le j\le K_1} |\gamma_2(i,j)|$ converges stochastically to zero as n increases.

This last expression is (2.61)* with $\Delta(6,n)$ given by

$$\Delta(6,n) = -\frac{\sqrt{m}}{L_2}\sum_{j=1}^{K_1}[Z_2(K_2+1,j)-Z_2(1,j)]$$

$$-\frac{1}{2L_2}\sum_{j=1}^{K_1}\left[\frac{g'_{2n}(\theta_2(K_2+1,j))Z_2^2(K_2+1,j)}{t_2^2(K_2+1,j)}-\frac{g'_{2n}(\theta_2(1,j),j)Z_2^2(1,j)}{t_2^2(1,j)}\right]$$

$$+(L_2-1)\sum_{j=1}^{K_1}\sum_{i=1}^{K_2}\left[-\frac{1}{8L_2^2}Q_2^2(i,j)-\frac{\sqrt{m}[Z_2(i+1,j)-Z_2(i,j)]}{2L_2^2}\right.$$
$$\left.+\frac{1}{3(1+\gamma_2(i,j))^3}\cdot\left[\frac{\sqrt{n}[Z_2(i+1,j)-Z_2(i,j)]}{L_2}+\frac{Q_2(i,j)}{2L_2}\right]^3\right]$$

Define $R_2(r,s)$ as

$$(L_2-1)\sum_{j=1}^{K_1}\sum_{i=1}^{K_2}\left[\left(\frac{\sqrt{n}[Z_2(i+1,j)-Z_2(i,j)]}{L_2}\right)^r\left(\frac{Q_2(i,j)}{L_2}\right)^s\right].$$

To show that $\Delta(6,n)$ converges stochastically to zero as n increases it is sufficient to show that $R_2(r,s)$ converges stochastically to zero as n increases for the following pairs (r,s): $(1,1),(0,2),(0,3),(1,2),(2,1),(3,0)$.

$$|R_2(r,s)| \le \frac{(L_2-1)D_2^s}{L_2^{r/2}b_2^{2s}L_2^s}\sum_{j=1}^{K_1}\sum_{i=1}^{K_2}\left[\left|\frac{Z_2(i+1,j)-Z_2(i,j)}{\sqrt{\frac{L_2}{m}}}\right|^r[Z_2^2(i+1,j)+Z_2^2(i,j)]^s\right]$$

and therefore

$$E(|R_2(r,s)|) \leq \frac{(L_2-1)D_2^s}{b^{2s}L_2^{s+(r/2)}} K_1K_2J_2(r,s) \leq \frac{mK_1D_2^s}{b^{2s}L_2^{s+(r/2)}} J_2(r,s)$$

where $J_2(r,s)$ is a finite number depending only on (r,s).

By (2.26) this last expression approaches zero as n increases and this implies that $R_2(r,s)$ converges stochastically to zero as n increases, for all the pairs (r,s) we are interested in and the proof of Lemma 2.8 is completed.

Lemma 2.9.

The expression (2.66) can be written as

$$\sum_{j=1}^{K_1+1} \log t_1(j) + \Delta(7,n) \tag{2.66*}$$

where $\Delta(7,n)$ converges stochastically to zero as n increases.

Proof of Lemma 2.9.

By the mean value theorem we can write

$$f_{1n}(T_1(j) + \frac{Z_1(j)}{\sqrt{n}t_1(j)}) = t_1(j) + \frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}t_1(j)}$$

for some $w'_1(j)$ in the open interval (B_{11},B_{12}) $(1 \leq j \leq K_1+1)$. Therefore (2.66) can be written as

$$\sum_{j=1}^{K_1+1} \log t_1(j) + \sum_{j=1}^{K_1+1} \log\left[1 + \frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}\, t_1^2(j)}\right]$$

and this expression is exactly (2.66)* where

$$\Delta(7,n) = \sum_{j=1}^{K_1+1} \log\left[1 + \frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}\, t_1^2(j)}\right] .$$

To show that $\Delta(7,n)$ converges stochastically to zero as n increases we are first going to show that

$$\max_{1\le j\le K_1+1} \left|\frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}\, t_1^2(j)}\right|$$

converges stochastically to zero as n increases. Since $w'_1(j)$ $(1\le j\le K_1+1)$ is in the open interval (B_{11},B_{12}) $|f'_{1n}(w'_1(j))| < D_1$ $(1\le j\le K_1+1)$ and by the definition of b_1 , $t_1(j) \ge b_1$ $(1\le j\le K_1+1)$ and therefore

$$\max_{1\le j\le K_1+1} \left|\frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}\, t_1^2(j)}\right| < \frac{D_1 \max_{1\le j\le K_1+1} |Z_1(j)|}{\sqrt{n}\, b_1^2} .$$

Using (2.51) and (2.52), and for any positive ε

$$P\left\{\frac{D_1 \max\limits_{1\le j\le K_1+1} |Z_1(j)|}{\sqrt{n}\, b_1^2} \le \varepsilon\right\} \ge 1 - 2\sum_{j=1}^{K_1+1} P\{\frac{Z_1(j)}{\sigma(Z_1(j))} > \frac{\sqrt{n}\, b_1^2\varepsilon}{D_1\sigma(Z_1(j))}\} \ge$$

$$\geq 1 - 2 \sum_{j=1}^{K_1+1} P\left\{\frac{Z_1(j)}{\sigma(Z_1(j))} > \frac{\sqrt{n}\, b_1^2 \epsilon}{D_1 \bar{\sigma}}\right\}$$

$$= 1 - 2(K_1+1)\left[1 - \Phi\left(\frac{\sqrt{n}\, b_1^2 \epsilon}{D_1 \bar{\sigma}}\right)\right]$$

$$> 1 - \frac{2(K_1+1)D_1\bar{\sigma}}{\sqrt{n}\, b_1^2 \epsilon}$$

and by (2.21) this last expression approaches one as n increases. Therefore we may write $\Delta(7,n)$ as

$$\sum_{j=1}^{K_1+1} [1 + \gamma_1^*(j)] \frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}\, t_1^2(j)}$$

where $\max_{1 \leq j \leq K_1+1} \gamma_1^*(j)$ converges stochastically to zero as n increases. We have

$$E\left\{\sum_{j=1}^{K_1+1} \left|\frac{f'_{1n}(w'_1(j))Z_1(j)}{\sqrt{n}\, t_1^2(j)}\right|\right\} \leq \frac{(K_1+1)D_1C_1}{\sqrt{n}\, b_1^2}$$

for some constant C_1. Using the expression (2.21) this completes the proof of Lemma 2.9.

<u>Lemma 2.10.</u>

The expression (2.67) can be written as

$$\sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1} \log t_2(i,j) + \Delta(8,n) \qquad (2.67)^*$$

where $\Delta(8,n)$ converges stochastically to zero as n increases.

<u>Proof of Lemma 2.10</u>.

By the mean value theorem we can write

$$g_{2n}(T_2(i,j) + \frac{Z_2(i,j)}{\sqrt{m}t_2(i,j)}, j) = t_2(i,j) + \frac{g'_{2n}(W'_2(i,j),j)Z_2(i,j)}{\sqrt{m}\; t_2(i,j)}$$

$$(1\le j\le K_1, 1\le i\le K_2+1)$$

for some $W'_2(i,j)$ in the open interval $(B_{21}(j),B_{22}(j))$ $(1\le j\le K_1)$.

Therefore (2.67) can be written as

$$\sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1} \log t_2(i,j) + \sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1} \log[1 + \frac{g'_{2n}(W'_2(i,j),j)Z_2(i,j)}{\sqrt{m}\; t_2^2(i,j)}]$$

and this expression is exactly $(2.67)^*$ where

$$\Delta(8,n) = \sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1} [1 + \frac{g'_{2n}(W'_2(i,j),j)Z_2(i,j)}{\sqrt{m}\; t_2^2(i,j)}] .$$

To show that $\Delta(8,n)$ converges stochastically to zero as n increases we are first going to show that

$$\max_{1\le j\le K_1, 1\le i\le K_2} \left|\frac{g'_{2n}(W'_2(i,j),j)Z_2(i,j)}{\sqrt{m}\; t_2^2(i,j)}\right|$$

converges stochastically to zero as n increases. Since $W_2'(i,j)$ is in the open interval $(B_{21}(j),B_{22}(j))$ then $|g_{2n}'(W_2'(i,j),j)| < D_2$ $(1 \le j \le K_1, 1 \le i \le K_2+1)$ and by definition of b_2, $t_2(i,j) \ge b_2$ $(1 \le j \le K_1, 1 \le i \le K_2+1)$ and therefore

$$\max_{1 \le j \le K_1, 1 \le i \le K_2+1} \left| \frac{g_{2n}'(W_2'(i,j),j)Z_2(i,j)}{\sqrt{m}\, t_2(i,j)} \right| \le \frac{D_2 \max\limits_{1 \le j \le K_1, 1 \le i \le K_2+1} |Z_2(i,j)|}{\sqrt{m}\, b_2^2} .$$

Using (2.51) and (2.52) and for any fixed positive ϵ

$$P\left\{ \frac{D_2 \max\limits_{i,j} |Z_2(i,j)|}{\sqrt{m}\, b_2^2} \le \epsilon \right\} \ge 1 - 2 \sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} P\left\{ \frac{Z_2(i,j)}{\sigma(Z_2(i,j))} > \frac{\sqrt{m}\, b_2^2 \epsilon}{D_2 \sigma(Z_2(i,j))} \right\}$$

$$\ge 1 - 2 \sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} P\left\{ \frac{Z_2(i,j)}{\sigma(Z_2(i,j))} > \frac{\sqrt{m}\, b_2^2 \epsilon}{D_2\, \bar{\sigma}} \right\}$$

$$= 1 - 2K_1(K_2+1)\left[1 - \Phi\left(\frac{\sqrt{m}\, b_2^2 \epsilon}{D_2\, \bar{\sigma}}\right)\right]$$

$$> 1 - \frac{2K_1(K_2+1)D_2\bar{\sigma}}{\sqrt{m}\, b_2^2 \epsilon}$$

and by expression (2.25), this last expression approaches one as n increases.

Then we may write $\Delta(8,n)$ as

$$\sum_{j=1}^{K_1} \sum_{i=1}^{K_2+1} [1+\gamma_2^*(i,j)] \frac{g_{2n}'(W_2'(i,j),j)Z_2(i,j)}{\sqrt{m}\, t_2^2(i,j)}$$

where $\max_{1\le j\le K_2+1,\, 1\le j\le K_1} |\gamma_2^*(i,j)|$ converges stochastically to zero as n increases. And since

$$E\left\{\sum_{j=1}^{K_1}\sum_{i=1}^{K_2+1}\left|\frac{g'_{2n}(W'_2(i,j),j)Z_2(i,j)}{\sqrt{m}\, t_2^2(i,j)}\right|\right\} \le \frac{K_1(K_2+1)D_2C_2}{\sqrt{m}\, b_2^2}$$

for some constant C_2, using the expression (2.25) this completes the proof of Lemma 2.10.

Lemma 2.11.

The expression (2.68) can be written as

$$-\frac{(K_1+1)}{2}\log 2\pi - K_1(L_1-\tfrac{1}{2})\log\frac{L_1}{n} - (np_1-1)\log p_1 - \frac{1}{2}\log p_1 \qquad (2.68)^*$$

$$- n(1-q_1)\log(1-q_1) - \frac{1}{2}\log(1-q_1) + \Delta(9,n)\ ;$$

where $\Delta(9,n)$ approaches zero as n increases.

Proof of Lemma 2.11.

In this proof we are going to use Stirling's formula,

$$[\text{for each positive integer } K,\ \log K! = \frac{1}{2}\log 2\pi \qquad (2.72)$$

$$+ (K+\tfrac{1}{2})\log K - K + \frac{a(K)}{K} \quad \text{where } |a(K)| < 1]$$

the expression (2.1) and the Taylor expansion of some functions, typified by $\log(L_1-1) = \log L_1 - \frac{1}{L_1} - \frac{1}{L_1^2}\ldots$. Then (2.68) can be written as

$$\frac{1}{2}\log 2\pi + (n+\tfrac{1}{2})\log n - n + \frac{a(n)}{n} - \frac{1}{2}\log 2\pi - (np_1 - \tfrac{1}{2})\log(np_1 - 1)$$

$$+ (np_1 - 1) - \frac{a(np_1-1)}{np_1-1} - \frac{1}{2}\log 2\pi - (n - nq_1 + \tfrac{1}{2})\log(n-nq_1)$$

$$+ (n-nq_1) - \frac{a(n-nq_1)}{n-nq_1} - \frac{K_1}{2}\log 2\pi - K_1(m+\tfrac{1}{2})\log m + K_1 m$$

$$- \frac{K_1 a(m)}{m} - \frac{(K_1+1)}{2}\log n$$

and recalling the definition of $m = L_1 - 1$, reordering and making some simplifications, this expression may be written as

$$- \frac{(K_1+1)}{2}\log 2\pi + K_1(L_1 - \tfrac{1}{2})\log n - K_1(L_1 - \tfrac{1}{2})\log L_1 - (np_1-1)\log p_1$$

$$- \frac{1}{2}\log p_1 - n(1-q_1)\log(1-q_1) - \frac{1}{2}\log(1-q_1) - 1 - K_1$$

$$- K_1(L_1 - \tfrac{1}{2})\log(1 - \frac{1}{L_1}) - (np_1 - \tfrac{1}{2})\log(1 - \frac{1}{np_1})$$

$$+ \frac{a(n)}{n} - \frac{a(np_1-1)}{np_1-1} - \frac{a(n-nq_1)}{n-nq_1} - \frac{K_1 a(L_1-1)}{L_1-1} \quad .$$

Therefore,

$$\Delta(9,n) = -1 - K_1 - K_1(L_1 - \tfrac{1}{2})\log(1 - \frac{1}{L_1}) - (np_1 - \tfrac{1}{2})\log(1 - \frac{1}{np_1}) + \frac{a(n)}{n}$$

$$- \frac{a(np_1-1)}{np_1-1} - \frac{a(n-nq_1)}{n-nq_1} - \frac{K_1 a(L_1-1)}{L_1-1}$$

and by assumptions (3) and (4), the expression (2.19) and (2.72), expanding the log's it follows that $\Delta(9,n)$ converges to zero as n increases.

QED

Lemma 2.12.

The expression (2.69) can be written as

$$K_1[-\frac{(K_2+1)}{2}\log 2\pi - K_2(L_2-\frac{1}{2})\log\frac{L_2}{m} - (mp_2-1)\log p_2 - \frac{1}{2}\log p_2$$

$$- [m(1-q_2)]\log(1-q_2) - \frac{1}{2}\log(1-q_2)] + \Delta(10,n) \qquad (2.69)^*$$

where $\Delta(10,n)$ approaches zero as n increases.

Proof of Lemma 2.12.

The proof of this lemma is straightforward using the same kind of argument as the proof of Lemma 2.11. (2.69) can be written as

$$K_1\left[\frac{1}{2}\log 2\pi + (m+\frac{1}{2})\log m - m + \frac{a(m)}{m} - \frac{1}{2}\log 2\pi - (mp_2-\frac{1}{2})\log(mp_2-1)\right.$$
$$+ (mp_2-1) - \frac{a(mp_2-1)}{mp_2-1} - \frac{1}{2}\log 2\pi - (m-mq_2+\frac{1}{2})\log(m-mq_2) + (m-mq_2)$$
$$- \frac{a(m-mq_2)}{m-mq_2} - \frac{K_2}{2}\log 2\pi - K_2(L_2-\frac{1}{2})\log(L_2-1) + K_2(L_2-1)$$
$$\left.- \frac{K_2a(L_2-1)}{L_2-1} - \frac{(K_2+1)}{2}\log m\ .\right]$$

Reordering, recalling expression (2.3) and making some simplifications this expression may be written as

$$K_1 \left[-\frac{(K_2+1)}{2} \log 2\pi + K_2(L_2-\tfrac{1}{2})\log m - K_2(L_2-\tfrac{1}{2})\log m - (mp_2-1)\log p_2 \right.$$
$$-\frac{1}{2}\log p_2 - m(1-q_2)\log(1-q_2) - \frac{1}{2}\log(1-q_2) - 1 - K_2$$
$$-K_2(L_2-\tfrac{1}{2})\log(1-\frac{1}{L_2}) - (mp_2-\tfrac{1}{2})\log(1-\frac{1}{mp_2}) + \frac{a(m)}{m} - \frac{a(mp_2-1)}{mp_2-1}$$
$$\left. -\frac{a(m-mq_2)}{m-mq_2} - \frac{K_2 a(L_2-1)}{L_2-1} \;. \right]$$

This expression is (2.69)* where $\Delta(10,n)$ is defined as

$$K_1[-1 - K_2 - K_2(L_2-\tfrac{1}{2})\log(1-\frac{1}{L_2}) - (mp_2-\tfrac{1}{2})\log(1-\frac{1}{mp_2}) + \frac{a(m)}{m}$$
$$-\frac{a(mp_2-1)}{mp_2-1} - \frac{a(m-mq_2)}{m-mq_2} - \frac{K_2 a(L_2-1)}{L_2-1}]$$

and by assumptions (8) and (9), expressions (2.23) and (2.72), expanding the log's it follows that $\Delta(10,n)$ converges to zero as n increases.

QED

Lemma 2.13.

The expression (2.70) can be written as

$$\frac{(K_1+1)}{2} \log 2\pi + \frac{1}{2} \log p_1 + \frac{1}{2} \log(1-q_1) + \frac{K_1}{2} \log \frac{L_1}{n} + \Delta(11,n) \qquad (2.70)^*$$

where $\Delta(11,n)$ approaches zero as n increases.

<u>Proof of Lemma 2.13.</u>

Define

$$\Delta(11,n) = -\frac{(K_1+1)}{2} \log \frac{nm}{L_1^2} - \frac{1}{2} \log\left[\frac{L_1^2 K_1}{n^2} + \frac{L_1(1-q_1+p_1)}{n}\right] - \frac{K_1}{2} \log \frac{L_1}{n} .$$

To complete the proof we must show that $\lim_{n\to\infty} \Delta(11,n) = 0$.
Simplifying, we may write

$$\Delta(11,n) = -\frac{(K_1+1)}{2} \log \frac{m}{L_1} - \frac{1}{2} \log \frac{n}{L_1} - \frac{1}{2} \log \frac{L_1^2 K_1}{n^2} - \frac{1}{2} \log\left(1 + \frac{n(1-q_1+p_1)}{K_1 L_1}\right)$$

and recalling the facts that $m = L_1 - 1$ and $K_1 L_1 = n(q_1 - p_1)$ it follows that $\Delta(11,n) = -[(K_1+1)/2]\log(1 - 1/L_1)$, and from expression (2.19) it follows that $\Delta(11,n)$ approaches zero as n increases.

QED

<u>Lemma 2.14.</u>

The expression (2.71) can be written as

$$\frac{K_1(K_2+1)}{2} \log 2\pi + \frac{K_1}{2} \log p_2 + \frac{K_1}{2} \log(1-q_2) + \frac{K_1 K_2}{2} \log \frac{L_2}{m} + \Delta(12,n) \qquad (2.71)^*$$

where $\Delta(12,n)$ approaches zero as n increases.

Proof of Lemma 2.14.

Define

$$\Delta(12,n) =$$

$$-\frac{K_1(K_2+1)}{2}\log\frac{m(L_2-1)}{L_2^2}-\frac{K_1}{2}\log\left[\frac{L_2^2K_2}{m^2}+\frac{L_2(1-q_2+p_2)}{m}\right]-\frac{K_1K_2}{2}\log\frac{L_2}{m}\,.$$

Simplifying, we may write

$$\Delta(12,n) = -\frac{K_1}{2}\log\frac{m}{L_2}-\frac{K_1(K_2+1)}{2}\log\left(1-\frac{1}{L_2}\right)-\frac{K_1}{2}\log\frac{L_2^2K_2}{m^2}$$

$$-\frac{K_1}{2}\log\left[1+\frac{m(1-q_2+p_2)}{K_2L_2}\right]$$

and recalling the fact that $K_2L_2 = m(q_2-p_2)$ it follows

$$\Delta(12,n) = -\frac{K_1(K_2+1)}{2}\log\left(1-\frac{1}{L_2}\right)$$

and from expression (2.23) it follows that $\Delta(12,n)$ approaches zero as n increases.

QED

Applying the results of Lemmas 2.3 through 2.14 to Lemma 2.2, it follows that with probability approaching one as n increases, we may write $\log (g(Z(n),n))/(h(Z(n),n))$ as the sum of expressions $(2.56)^*$, $(2.57)^*$, $(2.58)^*$, $(2.59)^*$, $(2.60)^*$, $(2.61)^*$, (2.62), (2.63), (2.64),

(2.65), $(2.66)^*$, $(2.67)^*$, $(2.68)^*$, $(2.69)^*$, $(2.70)^*$, and $(2.71)^*$. That is we may write $\log (g(Z(n),n))/(h(Z(n),n)) = \sum_{i=1}^{12} \Delta(i,n)$, and by the definitions of the $\Delta(\cdot,n)$ it follows that $\log (g(Z(n),n))/(h(Z(n),n))$ converges stochastically to zero as n increases and therefore the proof of Theorem 2.1 is completed.

Chapter III

A BIVARIATE TEST OF GOODNESS OF FIT

3.0. Introduction

In this chapter, we are going to develop a bivariate test of goodness of fit, for testing that the sample comes from a completely specified continuous bivariate distribution. The distribution theory on which the test is based was developed in Chapter II.

In Section 3.1, we will propose the test of goodness of fit based on the functions of order statistics defined in (2.7) and (2.8).

In Section 3.2, we show that in a particular case the random variable used as test criterion has under the alternative a known distribution and we give a concrete example, which satisfies, as n increases, all the assumptions of Section 2.2.

3.1. The General Case

Suppose that $(X_{1i}(n), X_{2i}(n))$ $(1 \leq i \leq n)$ are identically independently distributed random variables with common unknown probability density function $f(x_1, x_2, n)$ and we want to test the hypothesis that $f(x_1, x_2, n) = u(x_1, x_2)$ for all n, where $u(x_1, x_2)$ is a completely specified p.d.f. with corresponding c.d.f. $U(x_1, x_2)$. Denote by $u_1(x_1)$ the marginal p.d.f., by $U_1(x_1)$ the marginal c.d.f., $U_1^{-1}\left(\frac{np_1 + (j-1)L_1}{n}\right)$ by $T^*_1(j)$ $(1 \leq j \leq K_1 + 1)$, and $u_1(T^*_1(j))$ by

$t_1^*(j)$ $(1 \le j \le K_1+1)$ and let $W_1^*(j)$ be $\sqrt{n}\,t_1^*(j)[Y_1(np_1+(j-1)L_1)-T_1^*(j)]$ $(1 \le j \le K_1+1)$. Define $u_2(x_2,j)$ as the conditional marginal p.d.f., conditioned on the Y_1's, and $U_2(x_2,j)$ the corresponding c.d.f. $(1 \le j \le K_1)$, $T_2^*(i,j)$ as $U_2^{-1}\left(\frac{mp_2+(i-1)L_2}{m}, j\right)$ $(1 \le i \le K_2+1, 1 \le j \le K_1)$, and $t_2^*(i,j)$ as $u_2(T_2^*(i,j),j)$ $(1 \le i \le K_2+1, 1 \le j \le K_1)$. Let $W_2^*(i,j)$ be $\sqrt{m}\,t_2^*(i,j)[Y_2(mp_2+(i-1)L_2,j)-T_2^*(i,j)]$ $(1 \le i \le K_2+1, 1 \le j \le K_1)$. Define $V^*(n)$ as

$$\frac{nm}{L_1}\left[\frac{L_1[W_1^*(1)]^2}{np_1} + \frac{L_1[W_1^*(K_1+1)]^2}{n(1-q_1)} + \sum_{j=1}^{K_1}[W_1^*(j+1)-W_1^*(j)]^2\right]$$

$$+\frac{m(L_2-1)}{L_2^2}\sum_{j=1}^{K_1}\left[\frac{L_2[W_2^*(1,j)]^2}{mp_2} + \frac{L_2[W_2^*(K_2+1,j)]^2}{m(1-q_2)}\right.$$

$$\left.+\sum_{i=1}^{K_2}[W_2^*(i+1,j)-W_2^*(i,j)]^2\right].$$

Then the results above tell us that for all asymptotic purposes, when the hypothesis $f(x_1,x_2,n) = u(x_1,x_2)$ is true, $V^*(n)$ can be considered to have a chi-square distribution with d degrees of freedom.

We reject the hypothesis if $V^*(n)$ is "too large". Since $W_1^*(j) = \frac{t_1^*(j)W_1(j)}{t_1(j)} + \sqrt{n}\,t_1^*(j)[T_1(j)-T_1^*(j)]$ $(1 \le j \le K_1+1)$ and $W_2^*(i,j) = \frac{t_2^*(i,j)W_2(i,j)}{t_2(i,j)} + \sqrt{m}\,t_2^*(i,j)[T_2(i,j)-T_2^*(i,j)]$ $(1 \le i \le K_2+1, 1 \le j \le K_1)$ we can express $V^*(n)$ in terms of $W(n)$ and in many cases we can explicitly find the asymptotic distribution of $V^*(n)$ when the hypothesis is false. For such cases we can compute the asymptotic power of a test of goodness of fit which rejects the hypothesis when $V^*(n)$ is "too large".

3.2. A Particular Case

If under both the null hypothesis and the alternative X_1 and X_2 are independent random variables then $T_2^*(i,j)$, $T_2(i,j)$, $t_2^*(i,j)$, $t_2(i,j)$ $(1 \le i \le K_2+1, 1 \le j \le K_1)$, are real numbers (nonrandom quantities).

Then when the alternative is true, the distribution of $V^*(n)$ approaches the noncentral chi-square distribution with d-degrees of freedom and noncentrality parameter given by

$$\frac{nm}{L_1^2}\left\{\frac{L_1[t_1^*(1)]^2}{p_1}[T_1(1)-T_1^*(1)]^2+\frac{L_1[t_1^*(K_1+1)]^2}{1-q_1}[T_1(K_1+1)-T_1^*(K_1+1)]^2\right.$$
$$\left.+n\sum_{j=1}^{K_1}\Big[t_1^*(j+1)[T_1(j+1)-T_1^*(j+1)]-t_1^*(j)[T_1(j)-T_1^*(j)]\Big]^2\right\}$$
$$+\frac{m(L_2-1)}{L_2^2}\sum_{j=1}^{K_1}\left\{\frac{L_2[t_2^*(1,j)]^2}{p_2}[T_2(1,j)-T_2^*(1,j)]^2\right.$$
$$+\frac{L_2[t_2^*(K_2+1)]^2}{1-q_2}[T_2(K_2+1,j)-T_2^*(K_2+1,j)]^2$$
$$+m\sum_{i=1}^{K_2}\Big[t_2^*(i+1,j)[T_2(i+1,j)-T_2^*(i,j)]$$
$$\left.-t_2^*(i,j)[T_2(i,j)-T_2^*(i,j)]\Big]^2\right\}.$$

3.2.1. An Example

Consider the problem of testing the null hypothesis that $f(x_1,x_2,n)=\frac{1}{\pi^2[1+(x_1-\mu_1^*)^2][1+(x_2-\mu_2^*)^2]}$ versus the alternative hypothesis

that $f(x_1,x_2,n) = \dfrac{1}{\pi^2[1+(x_1-\mu_1)^2][1+(x_2-\mu_2)^2]}$. Since for both hypothesis, X_1 and X_2 are independent, we can easily compute the quantities defined above.

In this case $T_1^*(j) = \mu_1^* + \tan[\pi(p_1 + \frac{(j-1)L_1}{n} - \frac{1}{2})]$, $T_1(j) = T_1^*(j) - \mu_1^* + \mu_1$, $t_1^*(j) = \dfrac{\sin^2[\pi(p_1 + \frac{(j-1)L_1}{n})]}{\pi}$, $t_1(j) = t_1^*(j)$ $(1 \le j \le K_1)$ and $T_2^*(i,j) = \mu_2^* + \tan[\pi(p_2 + \frac{(i-1)L_2}{m} - \frac{1}{2})]$, $T_2(i,j) = T_2^*(i,j) - \mu_2^* + \mu_2$, $t_2^*(i,j) = \dfrac{\sin^2[\pi(p_2 + \frac{(i-1)L_2}{m})]}{\pi}$, $t_2(i,j) = t_2^*(i,j)$ $(1 \le i \le K_2+1, 1 \le j \le K_1)$.

We are going to show that there exist numbers p_1 , p_2 , q_1 , q_2 , L_1 , L_2 , K_1 , K_2 such that all the assumptions of Section 2.2 are satisfied for this example.

Let $L_1 = n^{\lambda_1}$, $m = L_1 - 1 = n^{\lambda_1} - 1$, $L_2 = m^{\lambda_2}$, $p_1 = 1 - q_1 = n^{\lambda_3}/n$, $p_2 = 1 - q_2 = m^{\lambda_4}/m$ and therefore $K_1 = n^{1-\lambda_1}(1 - \frac{2n^{\lambda_3}}{n})$, $K_2 = m^{1-\lambda_2}(1 - \frac{2m^{\lambda_4}}{m})$.

Assume that λ_1 , λ_2 , λ_3 , λ_4 are real numbers which satisfy the following assumptions.

$$0 < \lambda_1 < 1 \ , \ 0 < \lambda_2 < 1 \ , \ 0 < \lambda_3 < 1 \ , \ 0 < \lambda_4 < 1 \tag{3.0}$$

$$3 + \lambda_1 < 4\lambda_3 \tag{3.1}$$

$$5 < 4\lambda_3 + 1.5\lambda_1 \tag{3.2}$$

$$3 + \lambda_2 < 4\lambda_4 \tag{3.3}$$

$$5 < 4\lambda_4 + 1.5\lambda_2 \tag{3.4}$$

$$1 + 4\lambda_1 < 4\lambda_1\lambda_4 + 1.5\lambda_1\lambda_2 \tag{3.5}$$

By $a(3.x) + b(3.y)$ we mean multiply inequality (3.x) by a and (3.y) by b and add the resulting inequalities term by term. Then the inequalities given in Table II hold.

Table II

Number	Inequality	Implied By
(3.6)	$3 < 2\lambda_3 + 1.5\lambda_1$	$(3.2) + [-2<-2\lambda_3]$
(3.7)	$0.95 < \lambda_3$	$1.5(3.1) + (3.2)$
(3.8)	$6.5 < 6\lambda_3 + \lambda_1$	$\frac{10}{3}(3.7) + \frac{2}{3}(3.2)$
(3.9)	$3 < 2\lambda_4 + 1.5\lambda_2$	$(3.4) + [-2<-2\lambda_4]$
(3.10)	$0.95 < \lambda_4$	$1.5(3.3) + (3.4)$
(3.11)	$6.5 < 6\lambda_4 + \lambda_2$	$\frac{10}{3}(3.10) + \frac{2}{3}(3.4)$
(3.12)	$1 + 2\lambda_1 < 2\lambda_1\lambda_4 + 1.5\lambda_1\lambda_2$	$(3.5) + \lambda_1[-2<-2\lambda_4]$
(3.13)	$1 < 1.5\lambda_1$	(3.2) and $[\lambda_3<1]$
(3.14)	$1 < \lambda_1(1+\lambda_4)$	(3.13) and (3.10)

Since in this case $f_{1n}(x_1) = \frac{1}{\pi[1+(x_1-\mu_1)^2]}$ and $g_{2n}(x_2,j) = \frac{1}{\pi[1+(x_2-\mu_2)^2]}$ $(1 \le j \le K_1)$ then by definitions (2.11) through (2.18) it follows that

$$b_1 = \frac{\sin^2(\pi p_1)}{\pi}, \quad b_2 = \frac{\sin^2(\pi p_2)}{\pi}, \quad B_1 = B_2 = \frac{1}{\pi}, \quad B_{11} = -\infty,$$

$$B_{12} = +\infty, \quad B_{21}(j) = -\infty \text{ and } B_{22}(j) = +\infty \quad (1 \le j \le K_1) \tag{3.15}$$

Since $f'_{1n}(x)$ exists and $|f'_{1n}(x)| < 1$ for all x and all n, then assumption (12) holds and $D_1 = 1$ for all n.

Since $g'_{2n}(x,j)$ exists and $|g'_{2n}(x,j)| < 1$ for all x and all n, then assumption (13) holds and $D_2 = 1$ for all n.

Since by (3.0) and (3.15), $B_{11} = -\infty$, $B_{12} = +\infty$, $0 < p_1$, $q_1 < 1$ for all n and $t_1(1)$, $T_1(1)$, $t_1(K_1+1)$ and $T_1(K_1+1)$ are finite for all n, then assumptions (30) and (31) hold.

Since by (3.0) and (3.5), $B_{21}(j) = -\infty$, $B_{22}(j) = +\infty$ $(1 \le j \le K_1)$, $0 < p_2$, $q_2 < 1$ for all n and $t_2(1,j)$, $T_2(1,j)$, $t_2(K_2+1,j)$ and $T_2(K_2+1,j)$ $(1 \le j \le K_1)$ are finite for all n, then assumptions (32) and (33) hold.

With the Table III we are going to show that all the other assumptions hold.

Table III

#	Assumption	Condition For This Case	Restatement of Condition	Implied by
(1)	$\lim_{n\to\infty} p_1 = 0$	$\lim_{n\to\infty} n^{\lambda_3}/n = 0$	$\lambda_3 < 1$	(3.0)
(2)	$\lim_{n\to\infty} q_1 = 1$	$\lim_{n\to\infty} (1 - \frac{n^{\lambda_3}}{n}) = 1$	$\lambda_3 < 1$	(3.0)
(3)	$\lim_{n\to\infty} np_1 = \infty$	$\lim_{n\to\infty} n^{\lambda_3} = \infty$	$0 < \lambda_3$	(3.0)
(4)	$\lim_{n\to\infty} n(1-q_1) = \infty$	$\lim_{n\to\infty} n^{\lambda_3} = \infty$	$0 < \lambda_3$	(3.0)
(5)	$\lim_{n\to\infty} K_1 = \infty$	$\lim_{n\to\infty} n^{1-\lambda_1}(1 - \frac{2n^{\lambda_3}}{n}) = \infty$	$\lambda_1 < 1$, $\lambda_3 < 1$	(3.0)
(6)	$\lim_{n\to\infty} p_2 = 0$	$\lim_{n\to\infty} m^{\lambda_4}/m = 0$	$\lambda_4 < 1$	(3.0)
(7)	$\lim_{n\to\infty} q_2 = 1$	$\lim_{n\to\infty} (1 - \frac{m^{\lambda_4}}{m}) = 1$	$\lambda_4 < 1$	(3.0)
(8)	$\lim_{n\to\infty} \frac{K_1}{mp_2} = 0$	$\lim_{n\to\infty} \frac{n^{1-\lambda_1}(1-2n^{\lambda_3-1})(n^{\lambda_1}-1)}{(n^{\lambda_1}-1)(n^{\lambda_1}-1)^{\lambda_4}} = 0$	$1 < \lambda_1(1+\lambda_4)$	(3.14)
(9)	$\lim_{n\to\infty} \frac{K_1}{m(1-q_2)} = 0$	$\lim_{n\to\infty} \frac{n^{1-\lambda_1}(1-2n^{\lambda_3-1})(n^{\lambda_1}-1)}{(n^{\lambda_1}-1)(n^{\lambda_1}-1)^{\lambda_4}} = 0$	$1 < \lambda_1(1+\lambda_4)$	(3.14)

Table III (Continued)

#	Assumption	Condition For This Case	Restatement of Condition	Implied by
(10)	$\lim_{n\to\infty} K_2 = \infty$	$\lim_{n\to\infty} (n^{\lambda_1}-1)^{1-\lambda_2} \cdot \left[1 - \frac{2(n^{\lambda_1}-1)^{\lambda_4}}{(n^{\lambda_1}-1)}\right] = \infty$	$\lambda_2 < 1$	(3.0)
(14)	$b_1 \leq 1$ for all sufficiently large n	$\frac{\sin^2(\pi p_1)}{\pi} < 1$ for all n		(3.15)
(15)	$b_2 \leq 1$ for all sufficiently large n	$\frac{\sin^2(\pi p_2)}{\pi} < 1$ for all n		(3.15)
(16)	$\lim_{n\to\infty} \frac{L_1 D_1}{n b_1^2} = 0$	$\lim_{n\to\infty} \frac{n^{\lambda_1} n^4}{n n^{4\lambda_3}} = 0$	$3 + \lambda_1 < 4\lambda_3$	(3.1)
(17)	$\lim_{n\to\infty} \frac{B_1 K_1}{\sqrt{L_1}\, b_1} = 0$	$\lim_{n\to\infty} \frac{(n-2n^{\lambda_3})n^2}{n^{\lambda_1} n^{\lambda_1/2} n^{2\lambda_3}} = 0$	$3 < 2\lambda_3 + 1.5\lambda_1$	(3.6)
(17)	$\lim_{n\to\infty} \frac{B_1 K_1 D_1}{\sqrt{n}\, b_1^3} = 0$	$\lim_{n\to\infty} \frac{(n-2n^{\lambda_3})n^6}{n^{\lambda_1} n^{1/2} n^{6\lambda_3}} = 0$	$6.5 < 6\lambda_3 + \lambda_1$	(3.8)
(18)	$\lim_{n\to\infty} \frac{p_1 D_1^2}{n b_1^4} = 0$	$\lim_{n\to\infty} \frac{n^7}{n n^{7\lambda_3}} = 0$	$6 < 7\lambda_3$	(3.7)

Table III (Continued)

#	Assumption	Condition For This Case	Restatement of Condition	Implied by
(20)	$\lim\limits_{n\to\infty} \frac{(1-p_1)D_1^2}{nb_1^4} = 0$	$\lim\limits_{n\to\infty} \frac{n^7}{nn^{7\lambda_3}} = 0$	$6 < 7\lambda_3$	(3.8)
(21)	$\lim\limits_{n\to\infty} \frac{K_1\sqrt{D_1}}{b_1\sqrt{L_1}} = 0$	$\lim\limits_{n\to\infty} \frac{(n-2n^{\lambda_3})n^2}{n^{\lambda_1}n^{2\lambda_3}n^{\lambda_1/2}} = 0$	$3 < 2\lambda_3 + 1.5\lambda_1$	(3.6)
(22)	$\lim\limits_{n\to\infty} \frac{nD_1}{L_1^{1.5}b_1^2} = 0$	$\lim\limits_{n\to\infty} \frac{nn^4}{n^{1.5\lambda_1}n^{4\lambda_3}} = 0$	$5 < 4\lambda_3 + 1.5\lambda_1$	(3.2)
(23)	$\lim\limits_{n\to\infty} \frac{L_2D_2}{mb_2^2} = 0$	$\lim\limits_{n\to\infty} \frac{m^{\lambda_2}m^4}{mm^{4\lambda_4}} = 0$	$3 + \lambda_2 < 4\lambda_4$	(3.3)
(24)	$\lim\limits_{n\to\infty} \frac{B_2K_2}{\sqrt{L_2}b_2} = 0$	$\lim\limits_{n\to\infty} \frac{[m-2m^{\lambda_4}]m^2}{m^{\lambda_2}m^{\lambda_2/2}m^{2\lambda_4}} = 0$	$3 < 2\lambda_4 + 1.5\lambda_2$	(3.9)
(25)	$\lim\limits_{n\to\infty} \frac{B_2K_2D_2}{\sqrt{m}b_2^3} = 0$	$\lim\limits_{n\to\infty} \frac{[m-2m^{\lambda_4}]m^6}{m^{\lambda_2}m^{1/2}m^{6\lambda_4}} = 0$	$6.5 < 6\lambda_4 + \lambda_2$	(3.11)
(25)	$\lim\limits_{n\to\infty} \frac{P_2D_2^2}{mb_2^4} = 0$	$\lim\limits_{n\to\infty} \frac{m^6}{m^{7\lambda_4}} = 0$	$6 < 7\lambda_4$	(3.10)
(26)	$\lim\limits_{n\to\infty} \frac{(1-q_3)D_2^2}{mb_2^4} = 0$	$\lim\limits_{n\to\infty} \frac{m^6}{m^{7\lambda_4}} = 0$	$6 < 7\lambda_4$	(3.10)

Table III (Continued)

#	Assumption	Condition For This Case	Restatement of Condition	Implied by
(28)	$\lim\limits_{n\to\infty} \frac{K_1K_2\sqrt{D_2}}{\sqrt{L_2}b_2} = 0$	$\lim\limits_{n\to\infty} \frac{(n-2n^{\lambda_3})}{n^{\lambda_1}} \cdot \frac{[m-2m^{\lambda_4}]m^2}{m^{\lambda_2}m^{2\lambda_4}m^{\lambda_2/2}} = 0$	$1 + 2\lambda_1 < 2\lambda_4 + 1.5\lambda_1\lambda_2$	(3.12)
(29)	$\lim\limits_{n\to\infty} \frac{mK_1D_2}{L^{1.5}b_2^2} = 0$	$\lim\limits_{n\to\infty} \frac{(n-2n^{\lambda_3})m^4}{m^{1.5\lambda_2}m^{4\lambda_4}} = 0$	$1 + 4\lambda_1 < 4\lambda_1\lambda_4 + 1.5\lambda_1\lambda_2$	(3.5)

Therefore given numbers λ_1 , λ_2 , λ_3 , λ_4 which satisfy (3.0) through (3.5) all the assumptions from Section 2.2 hold.

Therefore since $\lambda_3 = \lambda_4 = 0.95$ and $\lambda_1 = \lambda_2 = 0.99$ satisfy (3.0) through (3.5), then $p_1 = 1 - q_1 = n^{-0.05}$, $L_1 = n^{0.95}$, $m = n^{0.99}-1$, $L_2 = m^{0.95}$, $p_2 = 1 - q_2 = m^{-0.05}$ satisfy all the assumptions, for this example. And therefore the results of Section 2.4 hold.

Then when the alternative is true $V^*(n)$, for large n , has approximately a noncentral chi-square distribution with d-degrees of freedom and noncentrality parameter given by

$$\frac{nm(\mu_1-\mu_1^*)^2}{L_1^2\pi^2}\left\{\frac{L_1\sin^4(\pi p_1)}{p_1}+\frac{L_1\sin^4(\pi q_1)}{1-q_1}\right.$$

$$\left.+\,n\sum_{j=1}^{K_1}\left[\sin^2[\pi(p_1+\frac{(j+1)L_1}{n})]-\sin^2[\pi(p_1+\frac{jL_1}{n})]\right]^2\right\}$$

$$+\frac{m(L_2-1)K_1(\mu_2-\mu_2^*)^2}{L_2^2\pi^2}\left\{\frac{L_2\sin^4(\pi p_2)}{p_2}+\frac{L_2\sin^4(\pi q_2)}{1-q_2}\right.$$

$$\left.+\,m\sum_{i=1}^{K_2}\left[\sin^2[\pi(p_2+\frac{(i+1)L_2}{m})]-\sin^2[\pi(p_2+\frac{iL_2}{m})]\right]^2\right\}\;.$$

Appendix A

ASYMPTOTIC EQUIVALENCE

A.1. Historical Background

The concept of asymptotic equivalence of two sequences of probability distributions, where all terms in both sequences were defined on the same space and there was no requirement on the existence or non-existence of a limiting distribution for both sequences, was introduced by Ikeda [9] as a way to formulate the problem of asymptotic approximation in a generalized form and to study asymptotic independence. He generalized a notion of asymptotic independence of two random variables introduced by Jeffroy [11] in the study of extreme values.

The notion was implicitly generalized by Weiss [20], studying asymptotic distribution theory for samples drawn from distributions approaching the uniform distribution as the sample size increases, where he considered two sequences of random variables where the n^{th} term of each sequence was defined on the same dimensional space, but the dimension increased with n ; and he gave a sufficient condition for the asymptotic equivalence of the two sequences as we will define it in this appendix. This result will be stated in this appendix as Theorem A.1.

Weiss [21] applied asymptotic equivalence again to prove the asymptotic joint normality of an increasing number of sample quantiles, where the distributions from which the samples were drawn were univariate continuous, defined on the interval $[0,1]$, the p.d.f.'s for all n were bounded from above and below by two constants independent of n , the absolute value of the second derivative was bounded by a constant

independent of n and the p.d.f.'s and its two first derivatives were continuous in the open interval $(0,1)$ and on the right at 0 and on the left at 1. In this case the n^{th} term of both sequences were defined in the $k(n)$-dimensional euclidean space with, $k(n) = n^{1-\delta}-1$, $0.75 < \delta < 1$.

Another application given by Weiss [22], used asymptotic equivalence to study the asymptotic joint normality of an increasing number of order statistics, where the distributions from which the samples were drawn were assumed to be univariate continuous, for each n, the derivative of the p.d.f. exists and its absolute value is less than a bound depending on n. The assumptions are that we move into both tails, but not too fast, and some compromises concerned with the speed in moving into the tail, the graduality of the increasing number of order statistics under consideration, the separation in order of those order statistics, and the smoothness of the p.d.f. It is also assumed that the order statistics are sample quantiles and that as n increases, the population tends to cover all the real line and there is a compromise between the speed at which it does it and all the other elements mentioned earlier. Another point is that Weiss, in the proof that for this case the sufficient condition (see Theorem A) holds, assumed that the actual distribution was the normal one. He didn't make use of this point in his earlier papers.

Ikeda and Matsunawa [10], generalizing Weiss [21] slightly, under essentially the same assumptions, formalized the concept of asymptotic equivalence for the case we are considering in this appendix.

Another application was given by Weiss [23] in studying the problem of testing the hypothesis that the sample came from the uniform

distribution, when the samples are drawn from a distribution with p.d.f. $1 + \frac{r(x)}{\sqrt{n}}$, x in the open interval $(0,1)$, $r(x)$ unknown but satisfying some regularity conditions. He proved the asymptotic normality of a gradually increasing number of order statistics, and their asymptotic sufficiency.

A.2. Asymptotic Equivalence of Sequences of Random Variables

Let $\{X^*(n)\}$, $\{Y^*(n)\}$ be sequences of random variables such that $X^*(n)$ and $Y^*(n)$ are $K(n)$ dimensional random variables, let $f_{X^*}(x^*(n))$ and $f_{Y^*}(y^*(n))$ denote the p.d.f. if they are continuous or its probability mass function if they are discrete. Let $\mathcal{C}$ be the collection of sequences of random variables such that the n^{th} term of the sequence is a $K(n)$ dimensional random variable, and let $\{R(n)\}$ be a sequence of measurable regions such that $R(n)$ is a $K(n)$ dimensional measurable region.

Definition A.1

The sequences $\{X^*(n)\}$, $\{Y^*(n)\}$ are asymptotically equivalent, (denoted by $\{X^*(n)\} \sim \{Y^*(n)\}$), if for any sequence $\{R(n)\}$

$$\lim_{n\to\infty} |P\{X^*(n) \in R(n)\} - P\{Y^*(n) \in R(n)\}| = 0 .$$

Weiss [20] and [21] has proved the following theorem.

Theorem A.1

If $\{X^*(n)\}$ and $\{Y^*(n)\}$ are elements of $\mathcal{C}$, a sufficient condition for $\{X^*(n)\} \sim \{Y^*(n)\}$ is that $\log f_{Y^*}(X^*(n))/f_{X^*}(X^*(n))$

converges stochastically to zero as n increases or equivalently that $f_{Y*}(X^*(n))/f_{X*}(X^*(n))$ converges stochastically to one as n increases.

Even though not needed for this thesis, it seems to us useful to introduce the following two lemmas.

<u>Lemma A.1</u>

Asymptotic equivalence is a equivalence relationship on $\mathcal{C}$.

<u>Proof of Lemma A.1</u>

From the definition A.1 it is obvious that for any $\{X^*(n)\} \varepsilon \mathcal{C}$ and any $\{Y^*(n)\} \varepsilon \mathcal{C}$

i) $\{X^*(n)\} \sim \{Y^*(n)\}$,

ii) $\{X^*(n)\} \sim \{Y^*(n)\}$ implies $\{Y^*(n)\} \sim \{X^*(n)\}$. Given $\{X^*(n)\}$, $\{Y^*(n)\}$, $\{Z^*(n)\}$ elements of $\mathcal{C}$ and any $\{R(n)\}$, from the fact that

$$
\begin{aligned}
&|P\{X^*(n) \varepsilon R(n)\}-P\{Z^*(n) \varepsilon R(n)\}| \\
&\quad \leq |P\{X^*(n) \varepsilon R(n)\}-P\{Y^*(n) \varepsilon R(n)\}| \\
&\quad + |P\{Y^*(n) \varepsilon R(n)\}-P\{Z^*(n) \varepsilon R(n)\}|
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\lim_{n\to\infty} |P\{X^*(n) \varepsilon R(n)\}-P\{Z^*(n) \varepsilon R(n)\}| \\
&\quad \leq \lim_{n\to\infty} |P\{X^*(n) \varepsilon R(n)\}-P\{Y^*(n) \varepsilon R(n)\}| \\
&\quad + \lim_{n\to\infty} |P\{Y^*(n) \varepsilon R(n)\}-P\{Z^*(n) \varepsilon R(n)\}|
\end{aligned}
$$

we get

iii) $\{X^*(n)\} \sim \{Y^*(n)\}$ and $\{Y^*(n)\} \sim \{Z^*(n)\}$ implies $\{X^*(n)\} \sim \{Z^*(n)\}$.

QED

Lemma A.2

If $\{X^*(n)\}$ and $\{Y^*(n)\}$ [illegible] elements of C and for every n , $X^*(n)$ is continuous $K(n)$-dimensional random variable and $Y^*(n)$ is a discrete $K(n)$-dimensional random variable then $\{X^*(n)\}$ and $\{Y^*(n)\}$ are not asymptotically equivalent.

Proof of Lemma A.2

For each n , let $R(n)$ be the $K(n)$-dimensional support of $Y^*(n)$. Then, for all n , $P\{X^*(n) \varepsilon R(n)\} = 0$ and $P\{Y^*(n) \varepsilon R(n)\} = 1$ and therefore the requirements of Definition A.1 are not fulfilled.

QED

BIBLIOGRAPHY

[1] Anscombe, F.J. (1963), "Tests of Goodness of Fit", Jour. Roy. Stat. Soc., Series B., Vol. 25, pp. 81-94.

[2] Bernoulli, D. (1735), "Recherches physiques et astronomiques...". Pièces qui on Remporté le Prix Double de l'Académie Royele des Sciences en 1734, Paris, pp. 93-122.

[3] Cochran, W.G. (1952), "The χ^2 Test of Goodness of Fit", Ann. Math. Stat., Vol. 23, pp. 315-345.

[4] Cramér, H. (1946), Mathematical Methods of Statistics, Princeton Univ. Press, p. 424.

[5] Feller, W. (1968), An Intoduction to Probability Theory and Its Applications, Vol. I, John Wiley & Sons, Inc.

[6] Fisher, R.A. (1924), "The conditions under which chi square measures the discrepancy between observation and hypothesis", Jour. Roy. Stat. Soc., Vol. 87, pp. 442-450.

[7] Gibbons, J.D. (1971), Nonparametric Statistical Inference, McGraw-Hill Book Co.

[8] Gumbel, E.J. (1943), "On the Reliability of the Classical χ^2 Test", Ann. Math. Stat., Vol. 14, pp. 253-263.

[9] Ikeda, S. (1963), "Asymptotic Equivalence of Probability Distributions with Applications to some Problems of Asymptotic Independence", Ann. Inst. Stat. Math., Vol. 15, pp. 87-116.

[10] Ikeda, S. and Matsunawa, T. (1972), "On the Uniform Asymptotic Joint Normality of Sample Quantiles", Ann. Inst. Stat. Math., Vol. 24, pp. 33-52.

[11] Jeffroy, J. (1959), "Contribution à la Théorie des Valuers Extrêmes", Publ. Inst. Univ. Paris, Vol. 8, pp. 123-184.

[12] Kolmogorov, A. (1933), "Sulla Determinazione Empirica di una Legge di Distribuzione", Giorn. Ist. Ital. Attuari, Vol. 4, pp. 1-11.

[13] Mann, H.B. and Wald, A. (1942), "On the Choice of the Number of Class Intervals in the Application of the Chi square Test", Ann. Math. Stat., Vol. 13, pp. 306-317.

[14] Massey, F.J., Jr. (1951), "The Kolmogorov-Smirnov Test for Goodness of Fit", Journ. Am. Stat. Assoc., Vol. 46, pp. 68-78.

[15] Neyman, J. (1949), "Contribution to the Theory of the χ^2 Test", Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, pp. 239-273.

[16] Pearson, K. (1900), "On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be Reasonably Supposed to have arisen from Random Sampling", Philos. Mag. Series 5, Vol. 50, pp. 157-172.

[17] Smirnov, H. (1939), "Sur les Écarts de la Courbe de Distribution Empirique", Recueil Mathématique (Matematiceskii Sbernik), N.S.6, pp. 3-26.

[18] Smirnov, N. (1948), "Table for Estimating the Goodness of Fit of Empirical Distributions", Ann. Math. Stat., Vol. 19, pp. 279-281.

[19] Yates, F. (1934), "Contingency Tables involving Small Numbers and the χ^2 Test", Jour. Roy. Stat. Soc. Suppl., Vol. 1, pp. 217-235.

[20] Weiss, L. (1965), "On Asymptotic Sampling Theory for Distributions Approaching the Uniform Distribution", Z. Wahrscheinlichkeitstheorie verw, Gerb 4, pp. 2.7-221.

[21] Weiss, L. (1969), "The Asymptotic Joint Distribution of an Increasing Number of Sample Quantiles", Ann. Inst. Stat. Math., Vol. 21, pp. 257-263.

[22] Weiss, L. (1973), "Statistical Procedures Based on a Gradually Increasing Number of Order Statistics", Communications in Statistics, Vol. 2(2), pp. 95-114.

[23] Weiss, L. (1974), "The Asymptotic Sufficiency of a Relatively Small Number of Order Statistics in Test of Fit", Annals of Statistics, Vol 2, pp. 795-802.